

OM of: US-09-294-298-4 to: GenEmbl.* out_format : pfs
 Date: Jan 19, 2001 12:44 AM
 About: Results were produced by the GenCore software, version 4.5,
 Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters: -MODE=frame+pz2.model -DEV=slp
 -Q=/cgn2.1/USPRO.spool/US09294298/runat_17012001_125501_27756/app_query.fasta_1.4359
 -DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
 -MINMATCH=0.100 -LOOPCL=0.000 -LOAPEXT=0.000 -QGAPOP=4.500
 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
 -EGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
 -DELEXT=7.000 -START=1 -MATRIX=blos62 -TRANS=human40.cdi
 -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
 -MAXLEN=2000000000 -USER=US09294298.@CGN1_13912 -NCPU=6
 -ICPU=3 -LONGLOG -NO_XLPXY -WAIT -THREADS=1

Search information block:
 Query: US-09-294-298-4
 Query length: 1376
 Database: GenEmbl.*
 Database sequences: 1118133
 Database length: -1736092196
 Search time (sec): 10440.680000

| score_list: | sfid | Orig | zScore | EScore | Len | Documentation |
|------------------|-----------|---------|----------|--------|-----|----------------------------------|
| gb_ro:AF058790 | + 7216.50 | 2903.19 | 1.5e-153 | 4140 | 1 | AF058790 Rattus norvegicus Syn |
| gb_ro:AF058789 | + 6894.00 | 2808.44 | 2.8e-148 | 4539 | 1 | AF058789 Rattus norvegicus Syn |
| gb_ro:AF058783 | + 6589.00 | 2774.21 | 2.3e-146 | 4265 | 1 | AF058783 Rattus norvegicus GTF |
| gb_ro:AF048976 | + 6415.00 | 2583.14 | 1.0e-135 | 4063 | 1 | AF048976 Rattus norvegicus syn |
| gb_ro:AF015962 | + 6308.50 | 2539.72 | 2.6e-133 | 4801 | 1 | AB015962 Rattus norvegicus mRNA |
| gb_hg20:AL161903 | + 4066.00 | 1625.80 | 2.1e-82 | 150956 | 1 | AL161903 Homo sapiens chromosome |
| gb_pr7:HSJ2570F3 | + 4060.50 | 1626.05 | 2.1e-82 | 94770 | 1 | AL050332 Homo sapiens chromo |
| gb_hg20:AL161906 | + 3987.00 | 1593.63 | 1.3e-80 | 169323 | 1 | AL161906 Homo sapiens chromo |
| gb_pr3:AF047711 | + 2603.00 | 1060.08 | 6.9e-51 | 4368 | 1 | AF047711 Homo sapiens ncgAP mRNA |
| gb_hg7:AC019800 | + 1263.50 | 511.55 | 2.5e-20 | 56366 | 1 | AC019800 Drosophila melanogaster |
| gb_hg4:AC012162 | + 1263.50 | 505.11 | 5.6e-20 | 191504 | 1 | AC012162 Drosophila melanogaster |
| gb_in1:AE003506 | + 1263.50 | 502.72 | 7.6e-20 | 300994 | 1 | AE003506 Drosophila melanogaster |
| gb_hg4:AC012161 | + 1257.00 | 502.41 | 7.9e-20 | 202741 | 1 | AC012161 Drosophila melanogaster |
| gb_in1:AB011280 | + 1172.00 | 488.97 | 4.4e-19 | 3984 | 1 | AB011280 Caenorhabditis elegans |
| gb_pr4:AK024488 | + 1103.50 | 461.22 | 1.6e-17 | 4287 | 1 | AK024488 Homo sapiens mRNA for |
| gb_pr4:AK022662 | + 1088.50 | 458.32 | 2.3e-17 | 2385 | 1 | AK022662 Homo sapiens cDNA FLJ |
| gb_hg6:AC016523 | + 1082.00 | 436.23 | 3.0e-16 | 65961 | 1 | AC016523 Drosophila melanogaster |
| gb_ro:AF053938 | + 1065.00 | 454.50 | 3.7e-17 | 829 | 1 | AF053938 Rattus norvegicus p133 |
| gb_hg22:AL365274 | + 710.00 | 284.58 | 1.1e-07 | 172027 | 1 | AL365274 Homo sapiens chromo |
| gb_in2:CEC0785 | + 676.50 | 278.93 | 2.2e-07 | 39660 | 1 | 246266 Caenorhabditis elegans |
| gb_pr6:HS593C16 | + 650.50 | 264.72 | 1.4e-06 | 81971 | 1 | AL035702 Human DNA sequence |
| gb_hg12:AC027051 | + 613.00 | 246.29 | 1.5e-05 | 157803 | 1 | AC027051 Homo sapiens chromo |
| gb_hg3:AC011492 | + 613.00 | 246.29 | 1.5e-05 | 157803 | 1 | AC011492 Homo sapiens chromo |
| gb_hg17:AC073765 | + 600.00 | 238.94 | 3.8e-05 | 237619 | 1 | AC073765 Mus musculus clone |
| gb_in1:AB011279 | + 502.00 | 228.29 | 0.0002 | 2276 | 1 | AB011279 Caenorhabditis elegans |
| gb_hg3:AC011703 | + 502.00 | 201.89 | 0.0043 | 159777 | 1 | AC011703 Drosophila melanogaster |
| gb_pr1:AB011110 | + 488.50 | 214.18 | 0.0009 | 5579 | 1 | AB011110 Homo sapiens mRNA for |
| gb_hg13:AC040900 | + 479.50 | 193.05 | 0.0135 | 155332 | 1 | AC040900 Homo sapiens chromo |
| gb_hg20:AL161906 | + 479.00 | 192.39 | 0.0147 | 169323 | 1 | AL161906 Homo sapiens chromo |
| gb_ro:RATGAPX | + 455.00 | 203.57 | 0.0035 | 3296 | 1 | L13151 Rat GTPase-activating p |
| gb_pr8:HUNGAB | + 437.50 | 196.33 | 0.0089 | 3456 | 1 | M23612 Human GTPase-activating |
| gb_pr8:HUNGABA | + 437.50 | 195.17 | 0.0103 | 4307 | 1 | M23379 Human GTPase-activating |
| gb_cm:BTGAPR | + 433.50 | 193.37 | 0.0130 | 4473 | 1 | X12602 Bovine mRNA for GTPase |
| gb_cm:BTG0857 | + 433.00 | 195.77 | 0.0095 | 2730 | 1 | U30857 Bos taurus R-Ras GTPase |
| gb_p12:BG272158 | + 430.50 | 194.03 | 0.0119 | 3144 | 1 | A272158 Blumeria graminis f. |
| gb_pat1:E13125 | + 429.50 | 194.54 | 0.0112 | 2647 | 1 | E13125 Bovine gene for R-Ras p |
| gb_in3:DM6012609 | + 420.00 | 189.27 | 0.0219 | 3502 | 1 | AJ012609 Drosophila melanogaster |
| gb_hg12:AC027051 | + 417.00 | 168.04 | 0.0341 | 156700 | 1 | AC027051 Homo sapiens chromo |
| gb_pr4:AK026441 | + 400.50 | 182.63 | 0.0514 | 2813 | 1 | AK026441 Homo sapiens cDNA: FLJ |
| gb_in1:AB011283 | + 396.50 | 187.93 | 0.0260 | 760 | 1 | AB011283 Caenorhabditis elegans |
| gb_ro:MMU20238 | + 387.00 | 175.25 | 0.1325 | 4107 | 1 | U20238 Mus musculus GTPase-acti |

gb_pr7:HSINSP4BP + 381.00 174.80 0.1404 2837 ! X89399 Homo sapiens mRNA fo
 gb_in1:AB011285 + 378.50 180.26 0.0697 833 ! AB011285 Caenorhabditis eleg
 gb_ro:AF086714 + 371.00 170.31 0.2497 3117 ! AF086714 Mus musculus rasga
 gb_pr3:AF086713 + 362.50 166.86 0.3885 3148 ! AF086713 Homo sapiens rasga

seq_name: gb_ro:AF058790

seq_documentation_block:
 LOCUS AF058790 4140 bp mRNA ROD 09-OCT-1998
 DEFINITION Rattus norvegicus SynGAP-b mRNA, complete cds.
 ACCESSION AF058790
 VERSION AF058790.1 GI:3722228
 KEYWORDS
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE 1 (bases 1 to 4140)
 AUTHORS Kim,J.H., Liao,D., Lau,L.F. and Haganir,R.L.
 TITLE SynGAP: a synaptic RasGAP that associates with the PSD-95/SAP90
 protein family
 JOURNAL Neuron 20 (4): 683-691 (1998)
 MEDLINE 88240917

REFERENCE 2 (bases 1 to 4140)
 AUTHORS Kim,J.H. and Haganir,R.L.
 TITLE Direct Submission
 JOURNAL Submitted (09-APR-1998) Neuroscience, Johns Hopkins Medical
 Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA

REFERENCE 3 (bases 1 to 4140)
 AUTHORS Kim,J.H. and Haganir,R.L.
 TITLE Direct Submission
 JOURNAL Submitted (08-OCT-1998) Neuroscience, Johns Hopkins Medical
 Institute, 725 N. Wolfe Street, PCTB 900, Baltimore, MD 21205, USA

REMARK
 COMMENT On Oct 9, 1998 this sequence version replaced gi:3065890.
 FEATURES
 Location/Qualifiers
 1..4140
 /organism="Rattus norvegicus"
 /db_xref="taxon:10116"
 /tissue_type="hippocampus"
 1..3750
 /note="neural RasGAP; contains PH domain, C2 domain, GAP
 domain in the N-terminal region, and a T/SXV motif at the
 C-terminus"
 /codon_start=1
 /product="SynGAP-b"
 /protein_id="AAC6351.1"
 /db_xref="GI:3722229"

CDS
 1..3750

/translation="MGLRPPTFTSGSGSGSLPPPSHROPLRRRCSSCCFPEYHLG
 RRRKSVGGKQYMEAPAPRPSQGLSRRLKSSIKRTKSPKLDRTSSRQLLP
 RFRSADHRAKLQMSKSHESLSPSSAAEALNLDSDIIKPVHSSILGQEPK
 FVTTSSGTGKFCACRAERKQWIKQAVKPNKSRVDNVLKMTIEARLPPK
 KRYCYELCDLMLVARTSKPRASGDTVEGHEFNPLPAVRLRLHLRDSKKR
 KDKAGYVGLVTPVATLAGRHTEOMYPTVPTGSGSGSGSGSGSGSGSGSGK
 KGCPAVRLKAYOTMPLMELYKEAEVYHYRMCLAVLEPALNVKGLKEVASAL
 VHLQSTKAKDQFLSDMASSVDREMERHLIPEHTLATKALEYMRLLGOKYKLA
 IGEFRLHTESENEVDPICTASSLAHQANLMCCALCKVNSHCYVFPRELKE
 VFASRLKCAEREDIDRLISASLEFLCPLMPSFLGMEQETPDQTSRTLIL
 IAKVNLANSKFTSKEDFLGFNFELEWMSQOFLYEINLDTLTNNSSEFYI
 DGLRSLTHALLVQLSKALLKLLPLRLSLDIALRNPNIQPSQSRSEA
 RSQPMVLRGPAEMOGYMRDLNSIDLQSFMRGLNSMDMARLSPYKEKPPPPP
 GGGKDLFYVSRPLPASVCTSSDITTEPKMLSVNKSVMQLDQPGGRLNS
 SVSNLAAGDGLLHSSQSLTAALGLRAPAGRLSGSGSITAAAGMRLSOMGVTTDG
 VPAQLRPLSFQNPFLHMAADGPPAGHSGSGSGHGGPPSHHHHHHHHRRGPEPP
 DTFAPHTGTSSEDSLSTGVPAAPASLIHSHISYSDERPGSGTDFTRQLSDQNLOHM
 LSPPTQITGPQAPRPGSGSGSGSGSGSGGGGGPPQQRKQSLTIVRAQRPSPSSG
 NLQSPPEYVSNPAPRQSLSKSGSGSGSGSGGGGGKLPSTTKHSHQTSQTLNPT
 MPASPTVAVSNMPLHSDTESAHIREEYKLKEYSKSMDSERLDRVKEYEIEHSL
 KRLHMSVEELRLDRHPMAEPLPEPKRLLDQAQSGSPFPVQYQTRV"
 902 a 1328 c 1182 g 728 t

BASE COUNT
 ORIGIN

alignment_scores:

Quality: 7216.50 Length: 1377
 Ratio: 5.245 Gaps: 1
 Percent similarity: 99.927 Percent identity: 99.927

alignment_block:

US-09-294-298-4 x AF058790 ..

Align seg 1/1 to: AF058790 from: 1 to: 4140

1 GlyProProArgProArgGlnGlyAlaProValGlnValProCysPr 17
 11 GGCCCCCCCCCGAGCCCGCTCAGGGGCTCCGGTTACAGGTTCCTGCC 60
 17 oLeuLeuProThrAlaSerLeuSerAlaAlaAlaAlaLeuProAlaAlaP 34
 61 CTTCTTCCACCGCCAGGCTCTCCGCCCGCTGCTCTTCTCTGCTGCTT 110
 34 heArgGlyAsnThrThrTrpValAlaArgGlyGlyArgValSerProGly 50
 111 TCCGGGGAATACCACTTGGTCGCTCAGGAGGAAGAGTCTC.CCGGG 159
 51 GlyLysGlnTyrSerMetGluAlaAlaProAlaAlaProPheArgProSe 67
 160 GGAACACGTACAGCATGAGCCCGCCCGCTGCGCCCTTCCGGCCCTC 209
 67 rGlnGlyPheLeuSerArgLeuLysSerSerileLysArgThrLys 84
 210 GCAAGGCTTCTCAGCCGAGGCTAAAGAGCTCCATCAAGGTACAAAGT 259
 84 erGlnProLysLeuAspArgThrSerPheArgGlnIleLeuProArg 100
 260 CACAACCCAACTTGACCGGACACAGCTTTCGACAGATCCTGCCCTGC 309
 101 PheArgSerAlaAspHisAspAlaArgLeuMetGlnSerPheLysG 117
 310 TTCGGAAGTGTGACCATGACCGGCCCGGCTGATGCAGAGCTTCAAAG 359
 117 uSerHisSerHisGluSerLeuLeuSerProSerSerAlaAlaAla 134
 360 GTCTCACTCCCATGAGTCTCTGCTGAGTCCACAGCAGTCTGCTGAGGCC 409
 134 euLluLeuAsnLeuAspGluAspSerileLysProValHisSerSer 150
 410 TGGAGCTCAACCTGGATGAAGACTCCATTAAGCCAGTACAGAGTCC 459
 151 IleLeuGlyGlnGluPheCysPheGluValThrThrSerSerGlyThrLy 167
 460 ATCTTGGCCAGGAGTCTGCTTTCAGGTAAACACATCGTCTGGACAA 509
 167 sCysPheAlaCysArgSerAlaAlaGluArgAspLysTrpIleGluAsnL 184
 510 ATGTTTGGCTGTGCTGCTGCGAGGAAAGGACAAATGGATTGAGAATC 559
 184 euGlnArgAlaValLysProAsnLysAspAsnSerArgArgValAspAsn 200
 560 TACAGAGGCTGTGAACCCCAACAAAGGACACAGCCCGGGGTAGATAAC 609
 201 ValLeuLysLeuTrpIleIleGluAlaArgGluLeuProProLysLysAr 217
 610 GTGCTGAACTATGGATCATAGAAGCTCGAGAGTGGCCCCCAAGAGCG 659
 217 gtyrtyrCysGluLeuLysLeuAspMetLeuTrpAlaArgThrThrS 234
 660 ATATTACTCGAGTTATGCTTGGAGACATGCTCTATGCACGAGCACTT 709
 234 erLysProArgSerAlaSerGlyAspThrValPheTrpGlyGluHisPhe 250
 710 CCAAGCCCGCTCAGGCTCAGAGACACTGCTTTTGGGGCGAGCACTTC 759
 251 GluPheAsnAsnLeuProAlaValArgAlaLeuArgLeuHisLeuTyrAr 267

760 GAGTTTAAACACCTGCTGCTGTCGGGGCTGCGCTGCATCTGTACCG 809
 267 gAspSerAspLysLysArgLysLysAspLysAlaGlyTyrValGlyLeuV 284
 810 TGACTCGACAAAAAGCGGAAGAGGACAAAGGACAGCTAGTTGGCCTGG 859
 284 alThrValProValAlaThrLeuAlaGlyArgHisPheThrGluGlnTrp 300
 860 TGACTGTTCCAGTGGCCACCTGCTGGGGCCACCTTCACAGAGCAGTGG 909
 301 TyrProValThrLeuProThrGlySerGlyGlySerGlyMetGlySe 317
 910 TACCCCGTACCCCTGCCAACAGGAAGTGGGGCTCTGGGGGTATGGGCTC 959
 317 rGlyGlyGlyGlySerGlyGlySerGlyGlySerGlyGlyLysGlyG 334
 960 GGGGGGAGGAGGGGGTTCAGGGGGCTCAGGGGGCAAGGAAAGAG 1009
 334 lycysProAlaValArgLeuLysAlaArgTyrGlnThrMetSerileLeu 350
 1010 GCTGTCCTGCTGTCGGCTGAAGCCCGCTTACCAGACAAATGAGTATCCTG 1059
 351 ProMetGluLeuTyrLysGluPheAlaGluTyrValThrAsnHisTyrAr 367
 1060 CCCATGAGCTATATAAGAGTTTGCAGAATATGTACCACCACTACCTG 1109
 367 gMetLeuCysAlaValLeuGluProAlaLeuAsnValLysGlyLysGluG 384
 1110 CATGCTGTGCTGCTGTCGGAGCCGCCCTCAATGTCAAGGGCAAGGAG 1159
 384 luValAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLysAlaLys 400
 1160 AGTGTGCTAGTGCCTGTTTCATCTGCAAGACACAGGCAAGGCCAAG 1209
 401 AspPheLeuSerAspMetAlaMetSerGluValAspArgPheMetGluAr 417
 1210 GACTTCCCTTTCAGACATGGCCATGTCCAGAGGTAGACCGTTCATGGAGCG 1259
 417 gLuhisLeuIlePheArgGluAsnThrLeuAlaThrLysAlaIleGluG 434
 1260 GGAACACCTCATATTCGCGCAGAACACGCTCGCCACTAAAGCCATAGAAG 1309
 434 luTyrMetArgLeuIleGlyGlnLysThrLeuLysAspAlaIleGlyGlu 450
 1310 AGTATATGAGACTGATTGGCCAGAAATACCTCAAGGATGCCATTGGGGAG 1359
 451 PheIleArgAlaLeuTyrGluSerGluGluAsnCysGluValAspProIl 467
 1360 TTCATCCGGGCTCTGTATGAATCTGAGGAGAACTGTGAAGTAGACCCCAT 1409
 467 eLysCysThrAlaSerSerLeuAlaGluHisGlnAlaAsnLeuArgMetC 484
 1410 CAAGTGCACAGCTCCAGTCTGGCAGAGCACCAGGCCAACCTCGCGGATGT 1459
 484 yscysGluLeuAlaLeuLysValValAsnSerHisCysValPhePro 500
 1460 GCTGTGAGTTGGCCCTGTGCAAGGTGTCAACTCCCATTTGCGTGTCCCG 1509
 501 ArgGluLeuLysGluValPheAlaSerTrpArgLeuArgCysAlaGluAr 517
 1510 AGGAGCTGAAGAGGTGTTTGCATCATGCGGGCTGCGCTGTGCAGAGCG 1559
 517 gGlyArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuA 534
 1560 GGGCCGGGAGGACATTCCTGACAGGCTGATCAGCGCTCGCTCTTCTCTGC 1609
 534 rgPheLeuLysProAlaIleMetSerProSerLeuPheGlyLeuMetGln 550
 1610 GCTTCTCTGCGCGCCATCATGTGCGCCAGTCTGTTGGACTGATGACG 1659
 551 GluTyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysVa 567

| | | |
|------|---|------|
| 1660 | GAGTACCAGATGAGCAGACCTCAGCAACCTCACCCCTCACTCCCAAGGT | 1709 |
| 567 | lileGlnAsnLeuAlaAsnPheSerLysPheThrSerLysGluAspPheL | 584 |
| 1710 | TATCAGAACTCGCCAACTTTTCCAAGTTTACTCTAAAGAGGAGCACTTC | 1759 |
| 584 | eUGlyPheMetAsnGluPheLeuGluLeuGluTrpGlySerMetGlnGln | 600 |
| 1760 | TGGGCTTCATGACGAGTTTCTGGAGCTGGAGTTGGGTTCTATGACAGCA | 1809 |
| 601 | PheLeuTyrgluIleSerAsnLeuAspThrLeuThrAsnSerSerSerPh | 617 |
| 1810 | TTCTTGTATGAGATATCCAACTCTGGACACACTGACCAACAGCAGCAGTTT | 1859 |
| 617 | eGluGlyTyrlleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuL | 634 |
| 1860 | TGAGGGCTACATAGACTTGGGCGCGAGCTCTCCACACTTCACGCCCTGC | 1909 |
| 634 | eUTrpGluValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGly | 650 |
| 1910 | TC TGAGGAGGTCTGCCCGAGCTCAGCAAGAGAGCCCTCTGAAGCTGGGC | 1959 |
| 651 | ProLeuProArgLeuLeuSerAspIleSerThrAlaLeuArgAsnProAs | 667 |
| 1960 | CGGCTGCCCGCGCTCCTCAGCGACATCAGCACAGCCCTGAGGAACCTAA | 2009 |
| 667 | nIleGlnArgGlnProSerArgGlnSerGluArgAlaArgSerGlnProM | 684 |
| 2010 | CATCCAAAGGCGAGCGGAGCGCCAGAGCGCGCTCGGCTCTCAGCCCA | 2059 |
| 684 | etValLeuArgGlyProSerAlaGluMetGlnGlyTyrlMetMetArgAsp | 700 |
| 2060 | TGGTCTCGCGGGCGGTCACCGCAATGCAAGGCTACATGATGCGGGAC | 2109 |
| 701 | LeuAsnSerSerIleAspLeuGlnSerPheMetAlaArgGlyLeuAsnSe | 717 |
| 2110 | CTCAACAGCTCCATGAGCTTCACTGCTTCATGCTCGAGGCCCTCAACAG | 2159 |
| 717 | rSerMetAspMetAlaArgLeuProSerProThrLysGluLysProProP | 734 |
| 2160 | CTCTATGACATGGCTCGCCTCCCTCCCCAACCAAGGAGAAACCCCCG | 2209 |
| 734 | roProProGlyGlyGlyLysAspLeuPheTyrlValSerArgProPro | 750 |
| 2210 | CGCCCTTCCCGTGGGGGTAAAGACCTTCTATGTAGCCGCGCCACCA | 2259 |
| 751 | LeuAlaArgSerSerProAlaTyrlCysThrSerSerSerAspIleThrGl | 767 |
| 2260 | CTGGCCCGGCTCTCCCCAGCATCTGCAACAGCAGCTCGGACATCACAGA | 2309 |
| 767 | uProGluGlnLysMetLeuSerValAsnLysSerValSerMetLeuAspL | 784 |
| 2310 | GC GGGAGCAGAGATGCTGAGTGTCACACAGAGTGTCTCCATGCTGGACC | 2359 |
| 784 | eUGlnGlyAspGlyProGlyGlyArgLeuAsnSerSerValSerAsn | 800 |
| 2360 | TGCAAGGGGACGGGCTCGGGGCGCCCTTAACAGACGAGTAGTGTTCACA | 2409 |
| 801 | LeuAlaAlaValGlyAspLeuLeuHisSerSerGlnAlaSerLeuThrAl | 817 |
| 2410 | CTGGCAGCTGTGGGAGCTGTTCATCTCAAGCCAGGCTTCATGACAGC | 2459 |
| 817 | alaLeuGlyLeuArgProAlaProAlaGlyArgLeuSerGlnGlySerG | 834 |
| 2460 | AGCCTTGGGGTTCGGGCTGCACTCGCGGGCGCCTCTCCCAAGGAGTAGT | 2509 |
| 834 | lySerSerIleThrAlaAlaGlyMetArgLeuSerGlnMetGlyValThr | 850 |
| 2510 | GCTCTTCCATCACAGCGCGGCATGCGCCTCAGCAGAGTGGGTGTCACT | 2559 |
| 851 | ThrAspGlyValProAlaGlnGlnLeuArgIleProLeuSerPheGlnAs | 867 |
| 2560 | ACGGATGTGTCCCGCCAGCAACTCGGCATCCCTCTTTCTTCTTCCAGAA | 2609 |

| | | |
|------|---|------|
| 867 | nProLeuPheHisMetAlaAlaAspGlyProGlyProAlaAlaGlyHisG | 884 |
| 2610 | CCCTCTCTTCATATGGCTGCCAGTGCACAGGCGCCCCCAGCAGGCCATG | 2659 |
| 884 | LyGlySerSerGlyHisGlyProProSerSerHisHisHisHisHisHis | 900 |
| 2660 | GAGGAGCAGTGGCCATGGTCCACCTTCCTCCCATCACACCACCACCAC | 2709 |
| 901 | HisHisHisHisArgGlyGlyGluProProGlyYAspThrPheAlaProPh | 917 |
| 2710 | CATCACCATCACCGAGGGGGAGAACCCCCAGGGGACACTTTTCCCGCGT | 2759 |
| 917 | eHisGlyTyrSerLysSerGluAspLeuSerThrGlyValProLysProp | 934 |
| 2760 | CCATGGCTATAGCAAGAGCGAGGACCTCTCTACAGGGGTCCCTTAAGCC | 2809 |
| 934 | roAlaAlaSerIleLeuHisSerHisSerTyrSerAspGluPheGlyPro | 950 |
| 2810 | CTGCGGGCTCCATCTCTACAGCACACAGCTACAGTGAATGAGTTTGACC | 2859 |
| 951 | SerGlyThrAspPheThrArgArgGlnLeuSerLeuGlnAspAsnLeuG1 | 967 |
| 2860 | TCTGGTACTGATTTTACCCTGCGCAGCTCTCACTTCAGGACAACTACA | 2909 |
| 967 | nHisMetLeuSerProGlnIleThrIleGlyProGlnArgProAlaP | 984 |
| 2910 | GCACATGCTCTCCCGCCCCAGATACCATCTGGTCCCCAGAGCCAGCTC | 2959 |
| 984 | roSerGlyProGlyGlySerGlyGlyGlySerGlyGlyGlyGlyGlyGly | 1000 |
| 2960 | CCTCAGGCCAGGAGGGGCGAGTGTGGGGCAGTGTGGGGCGGTGGG | 3009 |
| 1001 | GlyGlnProProLeuGlnArgGlyLysSerGlnGlnLeuThrValse | 1017 |
| 3010 | GGCCAGCCACTCCCTTCAGAGGGGAAATCTCAGCAGTGTACAGTGAG | 3059 |
| 1017 | rAlaAlaGlnLysProArgProSerSerGlyAsnLeuLeuGlnSerProG | 1034 |
| 3060 | TGCTGCCAGAAACCCCGCGCTCCAGCGGAACTATTGCAGTCCCGG | 3109 |
| 1034 | luProSerTyrGlyProAlaArgProArgGlnGlnSerLeuSerLysGlu | 1050 |
| 3110 | AACCAAGTTATGGTCTCCCTCCAGGCACACAGACCTTCAGCAAGAG | 3159 |
| 1051 | GlySerIleGlyGlySerGlyGlySerGlyGlyGlyGlyGlyGlyGly | 1067 |
| 3160 | GGCAGCATTTGGGGCAGCGGGGGCAGCGGTGGCGAGGGGTGGGGGCT | 3209 |
| 1067 | uLysProSerIleThrLysGlnHisSerGlnThrProSerThrLeuAsnP | 1084 |
| 3210 | CAAGCCCTCCATCACCAGCAGCATTCAGACCTCCATCCACGCTGAACC | 3259 |
| 1084 | roThrMetProAlaSerGluArgThrValIatrpValSerAsnMetPro | 1100 |
| 3260 | CCACATGTCGGGCTTCGAGGGGAGTGTAGTGGTGGGTCCAATATGCTC | 3309 |
| 1101 | HisLeuSerAlaAspIleGluSerAlaHisIleGluArgGluGluTyrLy | 1117 |
| 3310 | CACCTGTCCGTGACATCGAGGTGCACACATTGACGGGAGAGATACAA | 3359 |
| 1117 | tleuLysGluTyrSerLysSerMetAspGluSerArgLeuAspArgValL | 1134 |
| 3360 | GCTGAAGGAGTACTCGAAGTCCATGGACGAGAGCGGCTGTGACAGGTG | 3409 |
| 1134 | ysGluTyrGluGluGluIleHisSerLeuLysGluArgLeuHisMetSer | 1150 |
| 3410 | AGGAGTACGAGGAGAGATCCACTCACTGAGGAGAAAGGCTACACATG | 3459 |
| 1151 | AsnArgLysLeuGluGluTyrGluArgArgLeuLeuSerGlnGluGluG1 | 1167 |
| 3460 | ACCGGAGCTGGAAGACTACGAGCGGAGGCTGCTGTCCAGGAAGAGCA | 3509 |

used name: qb TO: AF058789

74 tgluLeuLysSerSerIleLysArgThrLysSerGlnProLysLeuAspArg 90
|||||
362 GCCTAAAGAAGCTCCATCAACAGTACAAAGTACACACCCAAAGCTTGACCGG 411
|||||
91 thrSerSerPheArgGlnIleLeuProArgPheArgSerAlaAspHisAs 107
|||||
412 ACCAGACAGCTTTCGACAGATCCCTGCCCTGCCCTCCGAAGTGTGACCATGA 461
|||||
107 pArgAlaArgLeuMetGlnSerPheLysGluSerHisSerHisGluSerL 124
|||||
462 CCGGCGCCGGCTGATGCAGAGCTTCAAGGAGTCTCACTCCCATGAGTCCC 511
|||||
124 euLeuSerProSerSerAlaAlaGluAlaLeuLeuLeuAsnLeuAspGlu 140
|||||
512 TGCTGAGTCCAGCAGTGTCTGTGAGGCCCTGGAGCTCAACCTGGATGAA 561
|||||
141 AspSerIleIleLysProValHisSerSerIleLeuGlyGlnGluPheCy 157
|||||
562 GACTCCATATCAAGCCAGTACACAGCTCCATCTGGCGCAGGAGTTCTG 611
|||||
157 sPheGluValThrThrSerSerGlyThrLysCysPheAlaCysArgSerA 174
|||||
612 CTTTGAGGTACACATCGTCTGGGACAAAATGTTTTGGCTGCTCGGTCTG 661
|||||
174 laAlaGluArgAspLysTrpIleGluAsnLeuGlnArgAlaValLysPro 190
|||||
662 CAGCCGAAAGGAGCAAAATGGATTGAGAAATCTACAGAGGGCTGTGAACCC 711
|||||
191 AsnLysAspAsnSerArgAlaGluValAspAsnValLeuLysLeuTrpIle 207
|||||
712 ACAAGGACACACAGCCGCGGTAGATAGCTGCTGAACATATGATCAT 761
|||||
207 eGluAlaArgGluLeuProLysLysArgTyrTyrCysGluLeuCysL 224
|||||
762 AGAAGCTCGAGAGTGCCTCCCAAGAGAGCATATTACTCGAGTTATGCC 811
|||||
224 euAspAspMetLeuTyrAlaArgThrThrSerLysProArgSerAlaSer 240
|||||
812 TGGAGCAGATGCTCTATGACGAGGACCACTTCCAAGCCCGCTCAGCCTCA 861
|||||
241 GlyAspThrValPheTrpGlyGluHisPheGluPheAsnAsnLeuProAl 257
|||||
862 GGAGACACTGCTTTTGGGGGGAGCACTTCGAGTTTAAACACTGCGCTGC 911
|||||
257 aValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysArgL 274
|||||
912 TGTCCGGCGCTGCGGCTGCATCTGTACCCTGACTCGGACAAAAGCGGA 961
|||||
274 yLysAspLysAlaGlyTyrValGlyLeuValThrValProValAlaThr 290
|||||
962 AGAAGGACAAAGCGAGGCTACCTGGCCCTGGTGTACTGTTCCAGTGGCCACC 1011
|||||
291 LeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeuProTh 307
|||||
1012 CTGGCTGGGCGCACTTCACAGACAGTGTGTACCCGTGACCTGACCTGCAAC 1061
|||||
307 rGlySerGlyLysSerGlyMetGlySerGlyGlyGlyGlyGlySerG 324
|||||
1062 AGAAGTGGGGCTCTGGGGGTATGGGCTCGGGGGAGGAGGGGGGTCTG 1111
|||||
324 yGlyGlySerGlyLysGlyLysGlyGlyCysProAlaValArgLeu 340
|||||
1112 GGGCGGGCTCAGGGGCAAGGGAAGAGGAGCTGTCTGCTGCTGCGGCTG 1161
|||||
341 LysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLysG 357
|||||
1162 AAGGCCGTTACAGACANTGAGTATCTCCCATGGAGCTATATNAGA 1211
|||||
357 uPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaValLeuG 374
|||||
1212 GTTTCAGAAATATGTGACAAACCACTACCGCATGCTGTGTGCGGTGCTGG 1261

374 luProAlaLeuAsnValLysGlyLysGluGluValAlaSerAlaLeuVal 390
|||||
1262 AGCCCGCCCTCAATGTCAAGGGCAAGGAGGAGTGCCTAGTGCACGTGTT 1311
|||||
391 HisIleLeuGlnSerThrGlyLysAlaLysAspPheLeuSerAspMetAl 407
|||||
1312 CACATCTCGAAAGCACAGCAAGGCCAAGGACTTCCCTTTAGACATGGC 1361
|||||
407 aMetSerGluValAspArgPheMetGluArgGluHisLeuIlePheArgG 424
|||||
1362 CATGTCAGAGGTAGACCGGTTTCATGGAGCGGGAACACCTCATATTCGCG 1411
|||||
424 luAsnThrLeuAlaThrLysAlaIleGluGluTyrMetArgLeuIleGly 440
|||||
1412 AGAACGCTCGCCACTAAAGCCATAGAAGATATATGAGACTGATGGC 1461
|||||
441 GlnLysTyrLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyrG 457
|||||
1462 CAGAAATACCTCAAGGATGCCATTGGGGAGTTTCATCCGGGCTCTGTATGA 1511
|||||
457 uSerGluGluAsnCysGluValAspProIleLysCysThrAlaSerSerL 474
|||||
1512 ATCTGAGGAGACTGTGAAGTAGACCCCATCAAGTGCACAGCGTCCAGTC 1561
|||||
474 euAlaGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCys 490
|||||
1562 TGGCAGAGCACCGAGGCCAACCCTGCGGATGTGCTGTGAGTTGGCCCTGTGC 1611
|||||
491 LysValValAsnSerHisCysValPheProArgGluLeuLysGluValPh 507
|||||
1612 AAGTGTGTCAACTCCCATTCGCTGCTCCCGAGGAGCTGAAGGAGGTCTT 1661
|||||
507 eAlaSerTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaA 524
|||||
1662 TGCAATATGCGGGCTGCGCTGTGCAGAGCGGGCGGGAGGACATGCTGT 1711
|||||
524 sPArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIle 540
|||||
1712 ACAGGCTGATCAGCGCTCGCTCTCTCTGCGCTTCTCTGCGCGGCATC 1761
|||||
541 MetSerProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGluTh 557
|||||
1762 ATGTGCGCCAGTCTGTTGGACTGTATGAGGAGTACCCAGATGACAGAC 1811
|||||
557 rSerArgThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnP 574
|||||
1812 CTCAGAACCCCTCACCTCATCCCAAGGTTATCCAGAACCTGGCGCACT 1861
|||||
574 heSerLysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPhe 590
|||||
1862 TTTCCAGTTTACCTCAAGAGGAGGACTTCTGGGCTTCATGAACGAGTTT 1911
|||||
591 LeuGluLeuGluTrpGlySerMetGlnGlnPheLeuTyrGluIleSerAs 607
|||||
1912 CTGGAGCTGGAGTGGGGTTCTATGCAGCAATCTTGTATGAGATATCCAA 1961
|||||
607 nLeuAspThrLeuThrAsnSerSerSerPheGluGlyTyrIleAspLeuG 624
|||||
1962 CCTGGACACACTGACCAACAGCAGCAGTTTTGAGGGCTACATAGACTTGG 2011
|||||
624 lyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProGln 640
|||||
2012 GCGCGGAGCTCTCCACACTTCAGCGCCCTGCTCTGGGAGGTGCTGCCCGC 2061
|||||
641 LeuSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSe 657
|||||
2062 CTCAGCAAGGAGCCCTCCTGAAGCTGGGCGCGCTGCCCGGCTCCTCAG 2111
|||||
657 rAspIleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnProSerA 674
|||||
2112 CGACATCAGCAGCCCTTGAGGAACCCCTAACATCCAAAGGCGAGCCGAGCC 2161
|||||
674 rGlnSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSer 690

```

|||||
2162 GCCAGAGGAGCGCGCTCGCTCAGCCCATGGTGTGCGGGCGGTCA 2211
691 AlaGluMetGlnGlyTyrMetMetArgAspLeuAsnSerSerIleAspLe 707
|||||
2212 GCCGAGATGCGAGGCGTACATGATGCGGGAGCTCAACAGCTCCATCGACCT 2261
707 uGlnSerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAlaArgL 724
|||||
2262 TCAGTCTCTTCACTGCTCGAGGCGCTCAACAGCTATATGGACATGGCTGCC 2311
724 euProSerProThrLysGluLysProProProProProGlyGlyGly 740
|||||
2312 TCCCTTCCCAACAGAGAAACCCCGCGGCCCTCCCGGTGGGGT 2361
741 LysAspLeuPheThrValSerArgProProLeuAlaArgSerSerProAl 757
|||||
2362 AAAGACCTGTTCTATGTAGCGCGGCCACCACTGGCCCGCTCCCTCCAGC 2411
757 aTyrCysThrSerSerSerAspIleThrGluProGluGlnLysMetLeuS 774
|||||
2412 ATACTGACGAGCAGCTCGACATCACAGCGCGGAGCAGAGATGCTGA 2461
774 erValAsnLysSerValSerMetLeuAspLeuGlnGlyAspGlyProGly 790
|||||
2462 GTGTCAACAAGAGTGTCTCATGCTGACCTGACGGCGACGGGCTGGG 2511
791 GlyArgLeuAsnSerSerSerValSerAsnLeuAlaValGlyAspLe 807
|||||
2512 GGCGCGCTTAAACAGCAGTGTGTCTTCAACCTGGCAGCTGTTGGGGAGCT 2561
807 uLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProA 824
|||||
2562 GTTGTACTCAAGCCAGCGCTTCACTGACAGCGCCITGGGGTTGCCGCTG 2611
824 IaProAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAla 840
|||||
2612 CACCTGCGCGCGCTCTCCCAAGGAGTGGCTCTTCCATCACAGCAGCC 2661
841 GlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaG 857
|||||
2662 GGCATGCGCCTCAGCCAGATGGTGTCTACGTACGGATGGTGTCCCGCCCA 2711
857 nGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaA 874
|||||
2712 GCACTGGCATCCCTCTTCTTCCAGAACCCCTCTTCCATATAGGCTG 2761
874 IaAspGlyProGlyProProAlaGlyHisGlySerSerGlyHisGly 890
|||||
2762 CCGATGGACACCGGCCCCCAGCAGGCGCATGGAGGGAGCAGTGGCCATGT 2811
891 ProProSerSerHisHisHisHisHisHisHisHisHisHisArgGlyG 907
|||||
2812 CCACCTTCTCCATACACACACACACACACATCACCATCCAGCGGGGG 2861
907 yGluProProGlyAspThrPheAlaProPheHisGlyTyrSerLysSerG 924
|||||
2862 AGAACCCCCAGGGAGACATTTTCCCGCGTTCCATGGCTATAGCAAGAGCG 2911
924 IaAspLeuSerThrGlyValProLysProProAlaAlaSerIleLeuHis 940
|||||
2912 AGGACCTCTTACAGGGGTCCTTAAGCCCTTCCGCTTCCATCTCTCAC 2961
941 SerHisSerThrSerAspGluPheGlyProSerGlyThrAspPheThrAr 957
|||||
2962 AGCCACAGCTACATGATGATTTGGACCTCTGTTACTGATTTTACCGG 3011
957 gArgGlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSerProProG 974
|||||
3012 TCGGCGAGCTCTCACTTCAGGACAACCTTACAGCAGATGCTCTCCCGGCC 3061
974 InIleThrIleGlyProGlnArgProAlaProSerGlyProGlyGly 990
|||||

```

```

3062 AGATCACCATCGGTCCAGAGGCCAGTCCCTCAGGGCCAGGAGGGGCG 3111
991 SerGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProProLeuG 1007
|||||
3112 AGTGTGGGGGCGAGTGGTGGGGGGGGGGGGCCAGCCACCTCCCTTGCA 3161
1007 nArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysProArgP 1024
|||||
3162 GAGGGGCAATCTCAGCAGTTGACAGTGTAGTGTGCTGCCAGAAACCCGCG 3211
1024 roSerSerGlyAsnLeuLeuGlnSerProGluProSerTyrGlyProAla 1040
|||||
3212 CGTCCAGCGGGAACCTATTCAGAGTCCCGGACCAAGTTATGGTCTGCC 3261
1041 ArgProArgGlnGlnSerLeuSerLysGluGlySerIleGlyLysSerG 1057
|||||
3262 CGTCCAGCGCAACAGAGCCCTCAGCAAGAGGGGAGCATTTGGGGGAGCGG 3311
1057 yGlySerGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleThrLysG 1074
|||||
3312 GSGCAGCGGTGGCGAGGGGTGGGGGCTCAAGCCCTCCATCACCAGC 3361
1074 InHisSerGlnThrProSerThrLeuAsnProThrMetProAlaSerGlu 1090
|||||
3362 AGCATTTCCAGACTCCATTCACGCTGAACCCACGATCGGGCTCGGAG 3411
1091 ArgThrValAlaIleTrpValSerAsnMetProHisLeuSerAlaAspIleG 1107
|||||
3412 CGGACTGTAGCTGGGTGCTCAATATATGCTCAGCTGTCCGCTGACATCGA 3461
1107 uSerAlaHisIleGluArgGluGluTyrLysLeuLysGluTyrSerLys 1124
|||||
3462 GAGTCCACACATTCAGCGGGAGAGTACAGCTGAAGGAGTACTCGAGT 3511
1124 erMetAspGluSerArgLeuAspArgValLysGlyTyrGluGluGluL 1140
|||||
3512 CCATGGAGAGAGCGCGCTGGACAGGTGAAGGAGTACAGGAGGAGATC 3561
1141 HisSerLeuLysGluArgLeuHisMetSerAsnArgLysLeuGluGly 1157
|||||
3562 CACTCAGTGAAGAAAGGCTACACATGTCCAAACCGGAAGCTGGAAGAGTA 3611
1157 xGluArgArgLeuLeuSerGlnGluGlnThrSerLysIleLeuMetG 1174
|||||
3612 CGAGCGGAGGCTGCTGCCAGGAGAGCAGACCAAGATCCTGATGC 3661
1174 InTyrGlnAlaArgLeuGluGlnSerGlyLysArgLeuArgGlnGln 1190
|||||
3662 AGTACCAAGCCCGCTGGAGCAGAGCAGAGAGCGCTTGAGGAGCAGCAG 3711
1191 ValGluLysAspSerGlnIleLysSerIleIleGlyArgLeuMetLeuVa 1207
|||||
3712 GTGGAGAGAGTCCAGATCAGAGCATCATTTGGCAGGCTGATGCTGT 3761
1207 lGluGluLeuArgArgAspHisProAlaMetAlaGluProLeuProG 1224
|||||
3762 GGAGAGAGAGTGGCGGGGACCAACCCGCCATGGCTGAGCGGCTGCCG 3811
1224 luProLysLysArgLeuLeuAspAlaGln..... 1233
3812 AACCCAAGAGAGGCTGCTCGAGCCTCAGCTCCTCATCAGGTAATTTCTCC 3861
1233 ..... 1233
3862 TGGTTCCGCTTTGACCAGGGGCGGAGACACTGGGGGAGGTGACTCCCGA 3911
1234 ..... ArgGlySerPheProProThrValGlnGlnThrArgVal. 1246
|||||
3912 CCACCTGCAGGAGAGGAGCTTCCCGCTTGGGTCCAAACAAACCCCGGTGT 3961
1247 ..ArgTrpProHisGlyThrAlaTrpProGlnProHisProPro 1262
|||||
3962 GACGCTGGCCCCACCTTTGGAACGGGCTGGCCCCCCCCAGCCCAACCCGCC 4011

```


477 CAATGATTGAGATCTACAGAGGGCTGTGAACCCCAACAGGACACA 526
 195 erArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGlu 211
 527 GCCCGGGGTAGATAACGTGCTGAACATATGATCATAGAAGCTCGAG 576
 212 LeuProProlLysArgTyrCysGluLeuCysLeuAspMetLe 228
 577 CTGCCCCCAAGAGGATATTACTGCGAGTTATCCCTGGAGCATGCT 626
 228 uTyrAlaArgThrSerLysProArgSerAlaSerGlyAspThrValP 245
 627 CTATGACAGGACACATTCACAGCCCGCTCAGCCTCAGGACACATGCT 676
 245 heTrpGlyGluHisPheGluPheAsnAsnLeuProAlaValArgAlaLeu 261
 677 TTGGGGCGGAGCACTTCGAGTTTAAACACCTGCCTGCTCCGGGCGCTG 726
 262 ArgLeuHisLeuTyrArgAspSerAspLysLysArgLysLysAspLysAl 278
 727 CGGCTGCATCTGTACCGTGACTCGGACAAAAGCGGAAGAGGACAAAGC 776
 278 aGlyTyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgH 295
 777 AGCTACGTGGCTGGTGCATGTTCCAGTGCCACCTGCCTGGCTGGGCGCC 826
 295 isPheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGly 311
 827 ACTTCACAGAGCAGTGGTACCCCGTGCACCTGCCAACAGGAAGTGGGGC 876
 312 SerGlyGlyMetGlySerGlyGlyGlyGlySerGlyGlySerGly 328
 877 TCTGGGGGTATGGCTCGGGGGAGAGGGGGTTCAGGGGGCGCTCAGG 926
 328 yGlyLysGlyLysGlyGlyCysProAlaValArgLeuLysAlaArgTyrG 345
 927 GGGCAAGGAAAGGAGGTGCTGCTGCTGCTGCGGCTGAAGCCCGCTTACC 976
 345 lnThrMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyr 361
 977 AGCAATGATATCTTGGCCATGGAGCTATATAAGGAGTTTGCAGAAATAT 1026
 362 ValThrAsnHisTyrArgMetLeuCysAlaValLeuGluProAlaLeuAs 378
 1027 GTACCAACACCTACCGCATGCTGTGTCGCTGCTGGAGCCCGCCCTCAA 1076
 378 nValLysGlyLysGluGluValAlaSerAlaLeuValHisIleLeuGlns 395
 1077 TGCAAGGGCAGGAGGAGGTGCTAGTGCACTGTTTCAATCTCTGCAAA 1126
 395 erThrGlyLysAlaLysAspPheLeuSerAspMetAlaMetSerGluVal 411
 1127 GCACAGCAAGGCCAAGGACTTCTTTTCAGACATGGCCATCTCAGAGGTA 1176
 412 AspArgPheMetGluArgGluHisIleLeuPheArgGluAsnThrLeuAl 428
 1177 GACCGGTTCATGAGGGGGAACCTCATATTCCCGAGAACACGCTCGC 1226
 428 aThrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuL 445
 1227 CACTAAAGCCATAGAAGATATATCAGACTGATTGGCCAGAAATACCTCA 1276
 445 ysAspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsn 461
 1277 AGGATGCCATTGGGAGTTTCATCCGGGCTCTGTATGAATCTTGAGGAGA 1326
 462 CysGluValAspProIleLysCysThrAlaSerSerLeuAlaGluHisGly 478
 1327 TGTGAAGTAGACCCATCAAGTGCACAGGCTCCAGTCTGGCAGAGACCA 1376
 478 nAlaAsnLeuArgMetCysGluLeuAlaLeuLeuCysLysValValAsnS 495
 1377 GGCAACCTGCGGATGTGTGAGTTGGCCCTGTGCAAGGTGTGCAACT 2326

495 erHisCysValPheProArgGluLeuLysGluValPheAlaSerTrpArg 511
 1427 CCCATTGCGTGTCCCGAGGAGCTGAAGGAGGTGTTGTCATCATGGCGG 1476
 512 LeuArgCysAlaGluArgGlyArgGluAspIleAlaAspArgLeuIleSe 528
 1477 CTGGCGTGTGCAGAGCGGGGCGGAGCATTTGCTGACAGGCTGATCAG 1526
 528 rAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSerProSerL 545
 1527 CGCCTCGCTCTTCTCGCTCTCTCTCGCCGCCATCATGTGCCCCAGTC 1576
 545 euPheGlyLeuMetGlnGluTyrProaspGluGlnThrSerArgThrLeu 561
 1577 TGTGTGGACTGATCAGAGGTACCCAGATGAGCAGACCTCAGAACCCCTC 1626
 562 ThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerLysPheTh 578
 1627 ACCCTCATCGCAAGGTATCCAGAACCTGSCCAACTTTTCCAAAGTTTAC 1676
 578 rSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGluLeuL 595
 1677 CTCAAAGGAGGACTTCTGGCTTCATGAACGAGTTTCTGGAGCTGGAAT 1726
 595 rpGlySerMetGlnGlnPheLeuTyrGluIleSerAsnLeuAspThrLeu 611
 1727 GGGTTCCTATGCAGCAATTTCTGTATGAGATATCCAACTGGACACACTG 1776
 612 ThrAsnSerSerPheGluGlyTyrIleAspLeuGlyArgGluLeuSe 628
 1777 ACCAACAGCAGCTTTTGAGGGCTATAGACTTGGCCGCGAGCTCTC 1826
 628 rThrLeuHisAlaLeuLeuTrpGluValLeuProGlnLeuSerLysGluA 645
 1827 CACACTTCACGCCCTGCTCTGGAGGTGCTGCCCCAGCTCAGCAAGGAAG 1876
 645 laLeuLeuLysLeuGlyProLeuProArgLeuLeuSerAspIleSerThr 661
 1877 CCCTCTGAAGCTGGGCCGCTGCCCGGCTCTCAGCGGCTCAGCATCAGACA 1926
 662 AlaLeuArgAsnProAsnIleGlnArgGlnProSerArgGlnSerGluAr 678
 1927 GCCCTGAGGAACCTTAACATCCAAAGGACGCGGAGCGCGAGAGCGCG 1976
 678 gAlaArgSerGlnProMetValLeuArgGlyProSerAlaGluMetGlnG 695
 1977 CGCTCGGCTCAGCCCATGCTGCTGCGGGGCGCTCAGCCGAATCGAGG 2026
 695 lyTyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerPheMet 711
 2027 GCTACATGATCGGGACCTCAACAGCTCCATCGACCTTCAGTCTTCAATG 2076
 712 AlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerProth 728
 2077 GCTCGAGCCCTCAACAGCTCTATGGACATGGCTCGCTCCCTCCCCCAAC 2126
 728 rLysGluLysProProProProGlyGlyGlyLysAspLeuPheTh 745
 2127 CAAGGAGAAACCCCGCCCGCTCCCGTGGGGTAAAGACCTGTTCT 2176
 745 yrValSerArgProProLeuAlaArgSerSerProAlaTyrCysThrSer 761
 2177 ATGTGAGCGGGCCACCTGGCCCGTCTCTCCCGAGACTGACGAGAGC 2226
 762 SerSerAspIleThrGluProGluGlnLysMetLeuSerValAsnLysSe 778
 2227 AGCTGGGACATCAGAGCGGAGCAGAGATGCTGAGTGTCAACAGAG 2276
 778 rValSerMetLeuAspLeuGlnGlyAspGlyProGlyArgLeuAsnS 795
 2277 TGTGTCCATGCTGGACCTGAGGGGCGAGGGGCTGGGGCGCGCTTAACA 2326

795 erSerSerValSerAsnLeuAlaValAlGlyAspLeuLeuHisSerSer 811
 2327 GCAGTAGTGTTCACACCTGGCAGCTGTGGGACCTGTGCTACTCAGC 2376
 812 GlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaProAlaGlyAr 828
 2377 CAGGCTTCACTGACAGAGCTTGGGGTGGGGCTGCACCTGCCGGGG 2426
 828 gLeuSerGlnGlySerGlySerSerIleThrAlaAlaGlyMetArgLeuS 845
 2427 CCTCTCCACAGGAGTGGCTCTTCCATCAGCAGCGGCGCTGCCTCA 2476
 845 erGlnMetGlyValThrThrAspGlyValProAlaGlnGlnLeuArgLle 861
 2477 GCCAGTGGGTGCTACCTACGGATGGGTGCCCGCCAGCAACTGCGCATC 2526
 862 ProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaAspGlyProGl 878
 2527 CCTCTTTCTTCCAGAACCTCTCTTCCATATGGCTGCCGATGACACAGG 2576
 878 yProProAlaGlyHisGlySerSerGlyHisGlyProProSerSerH 895
 2577 GCCCCACAGCGCCATGGAGGAGCAGTGGCCATGCTCCACCTTCCTCCC 2626
 895 isHisHisHisHisHisHisHisHisHisArgGlyGlyGluProProGly 911
 2627 ATCACCACCCACCCACCATCACCATCACCAGGGGGAGAACCCCGAGG 2676
 912 AspThrPheAlaProPheHisGlyTySerSerLysSerGluAspLeuSerTh 928
 2677 GACACTTTTGGCCGCTTCCATGGCTATAGCAAGAGCGAGGACCTCTCTAC 2726
 928 rGlyValProLysProProAlaAlaSerIleLeuHisSerHisSerTyrs 945
 2727 AGGGTCTCTAAGCCCTCTGGGCTCCATCTCTCAGACCCACAGCTACA 2776
 945 erAspGluPheGlyProSerGlyThrAspPheThrArgArgGlnLeuSer 961
 2777 GTGATGATTTGGACCTCTGTGACTGATTTTACCCCTGGCGACCTCTCA 2826
 962 LeuGlnAspAsnLeuGlnHisMetLeuSerProGlnIleThrIleG1 978
 2827 CTTACAGGACACCTACAGCATGCTCTCCCGCCGCCAGATCACCATCGG 2876
 978 yProGlnArgProAlaProSerGlyGlySerGlyGlyGlyS 995
 2877 TCCCCAGAGCCACTCTCTCAGGGCCAGAGGGGCGAGTGGTGGGGCA 2926
 995 erGlyGlyGlyGlyGlyGlyGlnProProProLeuGlnArgGlyLysSer 1011
 2927 GTGGTGGGGGGGTGGGGGCCAGCCACTCTCTTGCAGAGGGGCAAAATCT 2976
 1012 GlnGlnLeuThrValSerAlaAlaGlnLysProArgProSerSerGlyAs 1028
 2977 CAGCAGTGTACAGTGTGCTGCCAGAAACCCCGCCCTCCACCGGGA 3026
 1028 nLeuLeuGlnSerProGluProSerTyrglyProAlaArgProArgGlnG 1045
 3027 CCTATGTGAGTTCCTCCGGAACCAAGTATGGTCTGCTCCGCTCCAGCGAAC 3076
 1045 InSerLeuSerLysGluGlySerIleGlyGlySerGlyGlySerGlyGly 1061
 3077 AGAGCCTCAGCAAGAGGCGCAGCATTTGGGGGCGAGCGGGCGCGGTGGC 3126
 1062 GlyGlyGlyGlyGlyLeuLysProSerIleThrLysGlnHisSerGlnTh 1078
 3127 GGAGGGGTGGGGGCTCAAGCCCTCCATCACAAGACGACATTCACAGAC 3176
 1078 rProSerThrLeuAsnProThrMetProAlaSerGluArgThrValAlat 1095
 3177 TCCATCCACGCTGACCCACGATGCGGCTCGGAGCGGAGCTGATGCT 3226
 1095 rpValSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHisIle 1111

3227 GGGTGTCCATATGCTCCTACCTGTCCGCTGACATCGAGAGTGCACACTT 3276
 1112 GluArgGluGlyLysLeuLysGlyLysSerLysSerMetAspGlyLeu 1128
 3277 GAGCGGGAAGAGTACAAAGCTGAAGAGTACTCGAAGTCCATGGACGAGAG 3326
 1128 rArgLeuAspArgValLysGlyLysGlyLysGlyLysIleHisSerLeuLysG 1145
 3327 CCGACTGCACAGGTGAAGAGTACGAGGAGGAGATCCACTCACTGAAGG 3376
 1145 LuArgLeuHisMetSerAsnArgLysLeuGluGlyLysGlyLysArgLeu 1161
 3377 AAAGCTTACATGTCCAAACCGAAGCTGGAGAGTACGAGCGGAGGCTG 3426
 1162 LeuSerGlnGluGlnThrSerLysIleLeuMetGlnThrGlnAlaAr 1178
 3427 CTGTCCCGAGGAGGAGCAGCAGCAAGATCTGTGTCAGTACCAAGCCCG 3476
 1178 gLeuGluGlnSerGlyLysArgLeuArgGlnGlnValGlyLysAspS 1195
 3477 CCTGGACAGCGCAGAGCGCTTGGAGCAGCAGCGTGGAGAGGACT 3526
 1195 erGlnIleLysSerIleIleGlyArgLeuMetLeuValGluGluLeu 1211
 3527 CCCAGATCAAGAGCATCTTGGCAGCTGTGCTGGTGGAGGAGGACTG 3576
 1212 ArgArgAspHisProAlaMetAlaGluProLeuProGluProLysLysAr 1228
 3577 CGCGGGGACCCCGCCATGGCTGAGCGGTGCTGAACCCCAAGAGAG 3626
 1228 gLeuLeuAspAla. Gln... ArgGlySerPheProTrpValGlnGln 1243
 3627 GCTGCTCGACCTCAGTGGAGGAGGAGCTTCCCGCTTGGGTCCACAA 3676
 1244 ThrArgVal... ArgTrpProHisLeuGlyThrAlaTrpProGlnPr 1259
 3677 ACCCGCTGTGAGCTGGCGCCACCTTGGAGCGGCTGGCGCCCGCCAGCC 3726
 1259 oHisProProHisProGlyCysArgSerGlnArgThrAlaSerSerGlyT 1276
 3727 CCACCGCCCGCCACCGGCTGCAGATCAAGAGAACGGCGAGTTCCGGAA 3776
 1276 hrProGlnThrThrSerProProSerIleThrAspLeuLeuProCysAla 1292
 3777 CACCGAGACCACTAGCCACCCAGCATCACAGACCTCTTCCCTGTGCA 3826
 1293 ProTyrrProGlyProProSerValThrAspLeuLeuProSerAlaProAs 1309
 3827 CCTTACCGCCCGCCACCGCTGCAGCTCAGACCTCTTCCAGTGCACCGCA 3876
 1309 pProGlyThrSerProThrThrArgThrGlyArgHisGlnGlyThrAlaG 1326
 3877 CCCTGGAACATCAACCAACCCAGGACTGGACGTCCACCAAGGAGACGGG 3926
 1326 yLeuSerProLeuThrProProTrpGlyThrHisLeuSerThrProLeu 1342
 3927 GATGTCTCTTACCGCTCTTGGGGGCGCCATCTGTCAACCCCACTG 3976
 1343 LeuHisSerArgArgGluSerGlyThrLeuSerCysProLeuThrProGl 1359
 3977 TCTCATTCAGAGGAGGAGTGGGAGCCCTCAGCTGCCCTCTCACCCCGAG 4026
 1359 yHisHisLeuProHisThrAspProPheThrLeuGlyCysTyrrProHisP 1376
 4027 ACACCCACTACCCACACAGACCCCTTCACTCTGGGGTGTCTATCCCATC 4076
 1376 ro 1376
 4077 CT 4078

983 GGAGACACTGCTTTGGGGCGGAGCAGCTTCAGAGTTTAAACAACCTGCCTGC 1032
257 aValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLysArgL 274
|||||
1033 TGTCCGGGGCGTGGGGTGTGATCTGTACCGTGACTCGGACAAAAGCGGA 1082
274 yLysAspLysAlaGlyTyrValGlyLeuValThrValProValAlaThr 290
|||||
1083 AGAAGCAAGAGCAGGCTACGTTGGCTGTGTGACTGTTCCAGTGCGCAC 1132
291 LeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeuProTh 307
|||||
1133 CTGGCTGGGGCCACTTCACAGCAGAGTGATCCCGTGACCTGACCTGCCAAC 1182
307 rGlySerGlyGlySerGlyMetGlySerGlyGlyGlyGlyGlyGly 324
|||||
1183 AGGAAGTGGGGCTCTGGGGGTATGGGCTCGGGGGAGGAGGGGGTCTAG 1232
324 yGlyGlySerGlyGlyLysGlyLysGlyGlyCysProAlaValArgLeu 340
|||||
1233 GGGGGCGCTCAGGGGCAAGGAAAGAGGCTGTCTGCTGTGCGGCTG 1282
341 LysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyrLysG 357
|||||
1283 AAGGCCGCTTACCAACAATGAGTATCTGCCCATGGAGCTATATAAGGA 1332
357 uPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaValLeuG 374
|||||
1333 GTTTCAGAATATGTACCAACCACTACCGCATGCTGTGTGCGGTGCTGG 1382
374 LuProAlaLeuAsnValLysGlyLysGluGluValAlaSerAlaLeuVal 390
|||||
1383 AGCCCGCCCTCAATGTCAAGGGCAAGGAGGAGGTGCGTGTGACATGCTG 1432
391 HisIleLeuGlnSerThrGlyLysAlaLysAspPheLeuSerAspMetAl 407
|||||
1433 CACATCTCCAAAGCACAGGCAAGGCCAAGGACTTCTCTTCAGACATGCG 1482
407 eMetSerGluValAspArgPheMetGluArgGluHisLeuIlePheArgG 424
|||||
1483 CATGTCCAGAGGTAGACCGGTTTCATGGAGCGGGAACACCTCATATTCGGG 1532
424 LuAsnThrLeuAlaThrLysAlaIleGluGluTyrMetArgLeuIleGly 440
|||||
1533 AGAACACGCTCGCCACTAAAGCCATAGAAGAGTATATGACATGATGGC 1582
441 GlnLysTyrLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTyrG 457
|||||
1583 CAGAAATACCTCAAGGATGCCATTGGGGAGTTTCATCCGGGCTCTGTATGA 1632
457 userGluGluAsnCysGluValAspProIleLysCysThrAlaSerSerL 474
|||||
1633 ATCTGAGGAGAACTCTGAAGTAGACCCCATCAAGTGCACAGCCTCCAGTC 1682
474 euAlaGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCys 490
|||||
1683 TGCCAGAGCACCAGGCCAACCTCGCGGATGTCTGTGAGTTGGCCCTGTGC 1732
491 LysValValAsnSerHisCysValPheProArgGluLeuLysGluValph 507
|||||
1733 AAGGTGGTCACTCCCATGCGTGTCCCGAGGAGCTGAAGGAGGTGT 1782
507 eAlaSerTyrArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaA 524
|||||
1783 TGCATCATGGCGCTCGCTGTGCAGAGCGGGGCGGGGAGGACATTTGCTG 1832
524 spArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIle 540
|||||
1833 ACAGGCTGATACAGCCCTCGCTCTCTCTGCGCTCTCTCTCGCCGCGCATC 1882
541 MetSerProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnTh 557
|||||
1883 ATGTCGCCAGTCTGTTTGGACTGATGCAGGAGTACCCAGATGAGCAGAC 1932

557 rSerArgThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAlaAsnP 574
|||||
1933 CTCAGAACCCCTCACCTCATGCCAAGGTTATCCAGAACCTGGCCAAC 1982
574 heSerLysPheThrSerLysGluAspPheLeuGlyPheMetAsnGluPhe 590
|||||
1983 TTTTCCAGTTTACCTCAAGGAGGAGCTTCTTGGGCTTCATGAACGAGTTT 2032
591 LeuGluLeuGluTyrGlySerMetGlnGlnPheLeuTyrGluIleSerAs 607
|||||
2033 CTGGAGCTGGAGTGGGCTTCTATGCAGCAATTCCTGTA TGAGATATCAA 2082
607 nLeuAspThrLeuThrAsnSerSerSerPheGluGlyTyrIleAspLeuG 624
|||||
2083 CTTGACACACTGACCAACACAGCAGCTTTTGGGGCTACATAGACTTGG 2132
624 yArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProGln 640
|||||
2133 GCGCGAGCTCTCCACACTTCAGCGCTGCTCTGGAGGTGCTGCCCGAG 2182
641 LeuSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSe 657
|||||
2183 CTCAGCAAGGAGCCCTCTGAGCTGGCCCTGCCCCGGCTCTCTAG 2232
657 rAspIleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnProSerA 674
|||||
2233 CGACATCAGCAGCAGCCCTGAGGAACCTTAACATCCAAAGCGCCGAGCC 2282
674 rGlnSerGluArgAlaArgSerGlnProMetValLeuArgGlyProSer 690
|||||
2283 GCCAGAGCGAGCGGCTCGGCTCAGCCCATGCTGCTGGCGGGCGGTCA 2332
691 AlaGluMetGlnGlyTyrMetMetArgAspLeuAsnSerSerIleaspLe 707
|||||
2333 GCGGAGATGCGAGGCTTACATGATCGGGAGCTCAACAGCTCCATCGACCT 2382
707 uGlnSerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAlaArgL 724
|||||
2383 TCAGTCTCTTATGCTCGAGGCTCAACAGCTCTATGGACATGGCTCGCC 2432
724 euProSerProThrLysGluProProProProProProProGlyGlyGly 740
|||||
2433 TCCCTCCCAACCAAGGAGAAACCCCGCCGCCCTCCCGGTGGGGGT 2482
741 LysAspLeuPheTyrValSerArgProProLeuAlaArgSerSerProAl 757
|||||
2483 AAAGACCTGTTCTATGTGAGCGGCCACCACTGGCCCGGCTCTCCCCAGC 2532
757 atyrcysThrSerSerSerAspIleThrGluProGluGlnLysMetLeuS 774
|||||
2533 ATATGTCAGCAGCAGCTCGGACATCAGAGCGCGGAGCAGAGATGCTGA 2582
774 ervAlaLysLysSerValSerMetLeuAspLeuGlnGlyAspGlyProGly 790
|||||
2583 GTGTCAACAAGAGTGTCTCATGTGTCACCTGCGAGCGCGGCGGCGCTGG 2632
791 GlyArgLeuAsnSerSerValSerAsnLeuAlaAlaValGlyAspLe 807
|||||
2633 GCGCCCTTAACAGCAGTAGTGTTCACCACTGCCAGCTGTGGGACCT 2682
807 uLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProA 824
|||||
2683 GTTGACATCAAGCCAGGCTTCACTGACAGCAGGCTTGGGGTTCGGCCCTG 2732
824 laProAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAla 840
|||||
2733 CACCTGCCGGGCGCTCTCCCAAGGAGTGGCTCTTCCATCAGCAGCAGCC 2782
841 GlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaG 857
|||||
2783 GGCATGCGCCCTCAGCCAGATGGGTGTCTACTACGAGTGGTGTCTCCCGCCCA 2832

```

857 nGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaA 874
1174 InTySerGluAlaArgLeuGluGlnSerGluTySerGluArgGlnGln 1190
3783 AGTACCAAGCCCGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3832
1191 ValGluLysAspSerGlnIleLysSerIleIleGlyArgLeuMetLeuVa 1207
3833 GTGGAGAGAGACTCCAGATCAAGAGCATCATGGCAGGCTGATGCTGGT 3882
1207 IGluGluGluLeuArgAspHisProAlaMetAlaGluProLeuProG 1224
3883 GGAGGAGAGAGCTGCCCGGAGCACCACCCCGCATGGCTGAGCCGCTG 3932
1224 LuProLysArgLeuLeuAspAlaGlnArgGlySerPheProProTrp 1240
3933 AACCCCAAGAGAGGCTGCTCGACGCTCAGAGAGGAGCTTCCCGCTGG 3982
1241 ValGlnGlnThrArgVal...ArgTrpProHisLeuGlyThrAlaTrpPr 1256
3983 GTCCAAACAACCCCGCTGACGCTGCGCCACCTTGGAACGCGCTGGCC 4032
1256 OProGlnProHisProProHisProGlyCys 1266
4033 CCCCAGAGCCCGACCCCGCCCGCCCGCTGGC 4063

seq_name: gb_ro:AB016962

seq_documentation_block:
LOCUS AB016962 4801 bp mRNA ROD 01-SEP-2000
DEFINITION Rattus norvegicus mRNA for synGAP-d, complete cds.
ACCESSION AB016962
VERSION AB016962.1 GI:4417206
KEYWORDS synGAP-bi.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
REFERENCE 1 (bases 1 to 4801)
AUTHORS Suzuki,T.
TITLE SynGAP-d
JOURNAL Published Only in DataBase (1999) In press
REFERENCE 2 (bases 1 to 4801)
AUTHORS Suzuki,T.
TITLE Direct Submission
JOURNAL Submitted (14-AUG-1998) to the DBJ/EMBL/GenBank databases. Tatsuo Suzuki, Shinsu University School of Medicine, Department of Neuropathology; 3-1-1 Asahi, Matsumoto, Nagano 390-8621, Japan (E-mail:suzukit@sch.md.shinshu-u.ac.jp, Tel:+81-263-37-2683, Fax:+81-263-37-2725)
COMMENT On Mar 16, 1999 this sequence version replaced gi:4239945.
FEATURES
Location/Qualifiers
source 1..4801
/organism="Rattus norvegicus"
/sub_species="Sprague Dawley"
/db_xref="taxon:10116"
1124..4624
/gene="synGAP-d"
1124..4624
/gene="synGAP-d"
/feature="synGAP-d"
/codon_start=1
/product="synGAP-d"
/protein_id="BAA74972.1"
/db_xref="GI:4417207"

gene
CDS
/translaton="MEYFPAAPFRPSOGFLSRRLKSKTKSOPKLDRTSSPRLP
RFRSADHRLMQSFKESSHESLLSPSAAELNLDSESIKPVHSSILGQFPC
FYVTSSTGKCFACRAAKRWIENLQRAVKPNKRRVNDVNLKWIIEARLPPK
KRYTCELDDMLYARTTSPKRSASGDTFVGEHFEFNNLPVLRALRLYRSDKKR
KDKAGYVGLVTPVATLAGRHFTQWYPTLPTGSGSGGSGSGSGSGSGGKG

```

KGCPAVRLKARYQTMIIIPMELYKEFARYTNHYRMCLCAVLPEALNVKCKEVSASL
VHIQSTGKARDEFLSDMAMSEVDREMERHLLIFRENTLATRAIBEYMRLLGQKYLKDA
IGEFIRALYSEENCEVDPIKCTASLAEPHQANRLMCEALACKVYNHSHCVFPELKE
VFASWRLCAERGEREDIAURLISASLFLFLCPA LMSPLSFLGMOEY PDEQTSRTLT
IAKVIQNLANFSKFTSKEDFLGFMEFLGLEWGSQOFLYEISLNDITITNSSSPEGYI
DLGRLESTHALLAWELVLPOLSKIEALLKGLPLRLSLDSTALRNPNQOROPSPROSEBA
RSOPMVLGRPSAEQCYMMRDNLNSIDLOSEFARGLNSMDMARLPSPTEKPPPP
GGKDLFYVSRPPLARSSPAYCTSSSDITEPEQKMLSYNKSYSMLDLQDGGPGLNS
SSVSNLAAGDLHSSQASLITAAIGLRPAPAGRLSQSGSSITAGMKLSQWGYTTDG
VPAQDLRPLUSFNQPLFMAADGPPGAPGAGSSGHGPPSHHHHHHHHHRGGEPPG
DTFAPFGYKSDLSLSTGVKPPPAASILHSHSYSDEFGPSTDFTRQLSLQDNLQHM
LSPPQITIGQRAPSGPGGGGGGGGGGGPPPLQGRKSKQQLTVSAAQKPPSSG
NLQSPSPYGPAPRQOISLKESIGSGSGSGGGGGLKPSITKQHSOTPSLTNPT
MPASERTVAVSNMHLADIESAHTREYKLYKEYSKSMDSESLDRDEYEBEIHSLKE
PLHNSKLEERYERLLSQEQTSKILMYQARLBSQSEKRLRQQQVEKDSQIKSIIS
PSIQADNAGGGAAPGPRUG" 1043 a 1477 c 1314 g 967 t

BASE COUNT 1043 a 1477 c 1314 g 967 t

ORIGIN

alignment_scores:

Quality: 6308.50 Length: 1229
Ratio: 5.179 Gaps: 5
Percent Similarity: 99.105 Percent Identity: 99.024

alignment_block:

US-09-294-298-4 x AB016962 ..

Align seg 1/1 to: AB016962 from: 1 to: 4801

56 MetGluAlaLaProAlaLaProPheArgProSerGlnGlyPheLeuSe 72
1124 ATGAGATATTTTCCCGCTCGGCGCTTCGAGGCTTCCTGAG 1173
72 rArgArgLeuLysSerSerIleLysArgThrLysSerGlnProLysLeuA 89
1174 CCGAGGCTAAAAGCTCCATCAACGTACAAAGTCAACCAACCACTTG 1223
89 sPArgThrSerSerPheArgGlnIleLeuProArgPheArgSerAlaAsp 105
1224 ACCGGACACAGCTTTCGACAGATCTCGCTTCGCTCCGAGTCTGAC 1273
106 HisAspArgAlaArgLeuMetGlnSerPheLysGluSerHisSerHisG1 122
1274 CATGACGGCGCGCTGTATGACAGAGCTTCAAGAGTCTCACTCCCATGA 1323
122 uSerLeuLeuSerProSerSerAlaLaGluAlaLeuGluLeuAsnLeuA 139
1324 GTCCCTGCTGAGTCCAGCAGTGTCTGCTGAGGCGCTTGAGCTCAACCTGG 1373
139 spGluAspSerIleLysProValHisSerSerIleLeuGlyGlnGlu 155
1374 ATGAAGACTCCATTATCAAGCCAGTACACAGCTCCATCCTGGGCCAGGAG 1423
156 PheCysPheGluValThrThrSerSerGlyThrLysCysPheAlaCysAr 172
1424 TTCCTGCTTTCGAGTAAACAATCGTCTGGGACAAAATGTTTTCCTGTCG 1473
172 gSerAlaAlaGluArgAspLysTrpIleGluAsnLeuGlnArgAlaValL 189
1474 GTCTGCAGCCGAAAGGACAAATGGATTGAGAACTACAGAGGCGCTGGA 1523
189 ysProAsnLysAspAsnSerArgArgValAspAsnValLeuLysLeuTrp 205
1524 AACCCAAAGGACACAGCCCGGGGTAGATAACGTGCTGCTGAAACTATGG 1573
206 IleIleGluAlaArgGluLeuProProLysLysArgTyrCysGluLe 222
1574 ATCATAGAAGCTCGAGAGCTGCCCCCAAGAGCGATATTACTGCGAGTT 1623
222 uCysLeuAspAspMetLeuTyrAlaArgThrThrSerLysProArgSerA 239
1624 ATGCTGGAGGACATGCTCTATGACGGGACCTTCCAGCCCGCTCGAG 1673

239 laSerGlyAspThrValPheTrpGlyGluHisPheGluPheAsnLeu 255
1674 CCTCAGAGACACTGCTCTTTTGGGGCGAGCACTTCGAGTTTAAACAACCTG 1723
256 ProAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLy 272
1724 CCTGCTCTCCGGCGCTGCGGCTGCATCTGTACCTGACTCGGACAAAA 1773
272 sArgLysLysAspLysAlaGlyTyrValGlyLeuValThrValProValA 289
1774 GCGGAAGAAGGACCAAGGACAGGCTACGTTGGCTGGTGTGACTGTTCAGTGG 1823
289 laThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeu 305
1824 CCACCTTGGCTGGCGCCACTTTCACAGAGCAGTGTACCCCGTGACCCCTG 1873
306 ProThrGlySerGlyGlySerGlyMetGlySerGlyGlyGlyGlyGlyGly 322
1874 CCAACAGAGAGTGGGGCTCTGGGGGTATGGGCTCGGGGGGAGAGGGGG 1923
322 ySerGlyGlyGlySerGlyGlyLysGlyLysGlyGlyCysProAlaValA 339
1924 GTCAGGGGGCGCTCAGGGGGCAAGGAAAGAGAGGCTGCTCTGCTGTGC 1973
339 rgLeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeuTyr 355
1974 GGCTGAAGGCGCTTACCAGACAATGAGTATCTGCCCATGGAGCTATAT 2023
356 LysGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaVal 372
2024 AAGGAGTTTCAGAAATATGTGACCAACCCTACCGCATGCTGTGTGCGCT 2073
372 lLeuGluProAlaLeuAsnValLysGlyLysGlyGluValAlaSerAlaL 389
2074 GCTGGAGCGCGCTCAATGTCAAGGCAAGGAGGAGGCTCGCTAGTGAC 2123
389 euValHisIleLeuGlnSerThrGlyLysAlaLysAspPheLeuSerAsp 405
2124 TGGTTTCATCCTTCCAAAGCACAGGCAAGGCAAGGACTTCCTTTTCAGAC 2173
406 MetAlaMetSerGluValAspArgPheMetGluArgGluHisLeuIlePh 422
2174 ATGGCCATGTGAGAGGTAGACCGGTTTCATGAGGCGGGACACCTCATAT 2223
422 eArgGluAsnThrLeuAlaThrLysAlaIleGluGluTyrMetArgLeuI 439
2224 CCGCGAGACACGCTCCGCTCAAGCCATAGAAGAGATATATGAGACTGA 2273
439 leGlyGlnLysTyrLeuLysAspAlaIleGlyGluPheIleArgAlaLeu 455
2274 TTGGCCAGAAATACCTCAAGGATGCCATTTGGGAGTTTCATCCGGGCTCTG 2323
456 TyrGluSerGluGluAsnCysGluValAspProIleLysCysThrAlaSe 472
2324 TATGAATCTGAGGAGAACTGTGAAGTAGACCCCTCAAGTGCACACGCTC 2373
472 rSerLeuAlaGluHisGlnAlaAsnLeuArgMetCysGlyLeuAlaL 489
2374 CAGTCTGGCAGAGCACCAGGCCAACCTTGCCTGTGCTGTGAGTTGGCCC 2423
489 euCysLysValValAsnSerHisCysValPheProArgGluLeuLysGlu 505
2424 TGTGCAAGGTGGTCAACTCCCATTTGCTTCCGAGGAGGCTGAAGGAG 2473
506 ValPheAlaSerTrpArgLeuArgCysAlaGluArgGlyArgGluAsp1 522
2474 GTGTTTGCATCATGGCGCTGCGCTGTGACAGAGCGGGCGCGGAGGACAT 2523
522 eAlaAspArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCysProA 539
2524 TGCTGACAGGCTGATACGCGCTTCCTTCCTTCCTTCCTTCCTTCCTTC 2573

539 laIleMetSerProSerLeuPheGlyLeuMetGlnLysProAspGlu 555
556 GlnThrSerArgThrLeuThrLeuIleAlaLysValIleGlnAsnLeuAl 572
572 aAnPheSerLysPheThrSerLysGluAspPheLeuGlyPheMetAsnG 589
589 luPheLeuGluLeuGluTTPGlySerMetGlnGlnPheLeuTyrGluIle 605
605 2724 AGTTCTGGAGCTGGAGTGGGGTCTATGCAGCAATTCCTGTATGAGATA 2773
606 SerAsnLeuAspThrLeuThrAsnSerSerPheGluGlyTyrIleAs 622
622 2774 TCCAACCTGGACACACCTGACCAACAGCAGCTTTTGGAGGCTACATAGA 2823
622 pLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuP 639
639 2824 CTGGGGCCGGAGCTCTCACACTTCAGCCCTGCTGGAGGTGCTGC 2873
639 roGlnLeuSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeu 655
655 2874 CCAGCTCAAGCAAGGAGCCCTCCTGAAGCTGGGCCGCTGCCCGGGCTC 2923
656 LeuSerAspIleSerThrAlaLeuArgAsnProAsnIleGlnArgGlnPr 672
672 2924 CTCAGGACATCAGCACAGCCCTGAGAACCTTAACATCCAAAGCAGCC 2973
672 oSerArgGlnSerGluAlaArgSerGlnProMetValLeuArgGlyP 689
689 2974 GAGCCGCCAGAGCAGCGCGCTCGGTCTCAGCCCATGGTGTGCGCGGGC 3023
689 roSerAlaGluMetGlnGlyTyrMetMetArgAspLeuAsnSerSerIle 705
705 3024 CGTCAGCGAGATGACAGGCTACATGATGCGGGACCTCAACAGCTCCATC 3073
706 AspLeuGlnSerPheMetAlaArgGlyLeuAsnSerSerMetAspMetAl 722
722 3074 GACCTTCAGTCTTCATGCTCGAGCCCTCAACAGCTCTATGGACATGGC 3123
722 aArgLeuProSerProThrLysGluLysProProProProProProGlyG 739
739 3124 TCGCCCTCCCTCCCAACCAAGAGAGAAACCCCGCCGCCCTCCCGGTG 3173
739 lYcGlyLysAspLeuPheTyrValSerArgProProLeuAlaArgSerSer 755
755 3174 GGGGTAAAGACCTGTCTATGTAGCGCGGCCACCACTGGCCCGGTCTCC 3223
756 ProAlaTyrCysThrSerSerSerAspIleThrGluProGluGlnLysMe 772
772 3224 CCAGCATCTGCAAGAGAGCTCGGACATCACAGAGCGGAGCAAGAT 3273
772 tLeuSerValAsnLysSerValSerMetLeuAspLeuGlnGlyAspGlyP 789
789 3274 GCTGAGTGTCAACAGAGTGTCTCATGCTGACCTGACCTGAGGCGCAGGCG 3323
789 roGlyGlyArgLeuAsnSerSerValSerAsnLeuAlaAlaValGly 805
805 3324 CTGGGGCGCGCTTAACAGCAGTAGTGTTCACCTGCGAGCTGTTGGG 3373
806 AspLeuLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuAr 822
822 3374 GACCTGTGCTCAAGCAGGCTCTCCTGACAGCAGGCTTTGGGGTGGC 3423
822 gProAlaProAlaGlyArgLeuSerGlnGlySerGlySerIleThrAr 839
839 3424 GCCTGACCTGCGGGCGGCTCTCCAAAGGAGTGGCTCTTCCATCAG 3473
839 laAlaGlyMetArgLeuSerGlnMetGlyValThrAspGlyValPro 855

3474 CAGCGCGCATCGCGCTCAGCAGATGGTGTCTACTACGATGGTGTCCCC 3523
856 AlaGlnGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMe 872
872 3524 GCCAGCAGCTGCGCATCCCTCTTCTTCCATCCAGAACCTCTCTTCCATAT 3573
872 tAlaAlaAspGlyProGlyProProAlaGlyHisGlyGlySerSerGlyH 889
889 3574 GGCTGCCGATGGACAGCGGCCCCAGCAGGCCATGGAGGAGCAGTGGCC 3623
889 lSgLYProProSerSerHisHisHisHisHisHisHisHisHisHisArg 905
905 3624 ATGGTCCACCTCTCTCCATCCACCCACCCACCATCACCATCACCACCA 3673
906 GlyGlyGluProProGlyAspThrPheAlaProPheHisGlyTyrSerLy 922
922 3674 GGGGAGAAACCCAGGCGGACACTTTTGGCCCGTTCATGGCTATAGCAA 3723
922 sSerGluAspLeuSerThrGlyValProLysProProAlaAlaSerIleL 939
939 3724 GAGCGAGACCTCTCTACAGGGTCCCTAAGCCCTCGGCCCTCCATCC 3773
939 euHisSerHisSerTyrSerAspGluPheGlyProSerGlyThrAspPhe 955
955 3774 TTCAGCGACAGCTACAGTATGAGTTTGGACCTCTGGTACTGATTTT 3823
956 ThrArgArgGlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSerPr 972
972 3824 ACCGTCGGCAGCTCTCCTTCAGGACAACTACAGCACATGCTCTCCCC 3873
972 oProGlnIleThrIleGlyProGlnArgProAlaProSerGlyProGlyG 989
989 3874 GCCCGAGATCACCATCGGTCGCCAGAGCCAGCTCCCTCAGGGCCAGGAG 3923
989 lYcLYSerGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProPro 1005
1005 3924 GGGCAGTGGTGGGGCAGTGGTGGGGCGGTGGGGCCAGCCACCTCC 3973
1006 LeuGlnArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLysPr 1022
1022 3974 TTGCAGAGGGGCAATCTCAGAGTTCACAGTGTGCTGCTGCCAGAACCC 4023
1022 oArgProSerSerGlyAsnLeuLeuGlnSerProGluProSerTyrGlyP 1039
1039 4024 CCGGCCCTCCAGCGGAACTATTGACGTCGCCGGAACCAAGTTATGGTC 4073
1039 toAlaArgProArgGlnLeuSerLeuSerLysGlySerIleGlyGly 1055
1055 4074 CTGCCCGTCCACGCAACAGAGCTCAGCAAGAGAGGCGCATTTGGGGC 4123
1056 SerGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleTh 1072
1072 4124 AGCGGGGCGAGCGGTGGCGGAGGGGTGGGGGCTCAAGCCCTCCCATCAC 4173
1072 rLysGlnHisSerGlnThrProSerThrLeuAsnProThrMetProAlaAs 1089
1089 4174 CAAGCAGATTCACAGACTCCTCATCCAGCTGAACCCACGATGCCGCTC 4223
1089 erGluArgThrValAlaIleValSerAsnMetProHisLeuSerAlaAsp 1105
1105 4224 CGAGCGGACTGTAGCTGGGTGTCCAAATATGCTCACCTGTCCGCTGAC 4273
1106 lLeGluSerAlaHisIleGluArgGluGluTyrLysLeuLysGluTyrSe 1122
1122 4274 ATCAGAGTGTACACATTTGACGCGGAAAGATACAGCTGAAGGAGTACTC 4323
1122 rLysSerMetAspGluSerArgLeuAspArgValLysGluTyrGluGluG 1139
1139 4324 GAAGTCCATGACAGAGCGGCTGACAGG.....GAGTACGAGGAGG 4367
1139 luIleHisSerLeuLysGluArgLeuHisMetSerAsnArgLysLeuGlu 1155


```

4368 AGATCCACTACTGAAGAAAGGCTACACATGTCCAAACCGAAGCTGAA 4417
1156 GluTyrGluArgArgLeuLeuSerGlnGluGlnThrSerLysIleLe 1172
|||||
4418 GAGTACAGGCGAGGCTGTCTCCAGAGAGACGACGACGAGATCCT 4467
1172 uMetGlnTyrGlnAlaArgLeuGluGlnSerGluLysArgLeuArgGlnG 1189
|||||
4468 GATGACGATACCAAGCCCGCTGGAGCAGACGAGAGCGCTTGAGCGACG 4517
1189 InGlnValGluLysAspSerGlnIleLysSerIleIleGly..... 1202
|||||
4518 AGCAGGTGGAGAGGACTCCACATCAAGAGCATCATTTGGCAGCCGCTCC 4567
1203 ...Arg.LeuMetLeuValGluGluGluLeuArgArgAspHisProAlaM 1218
4568 CTTTCAAGGCTGATGTGTGTGAGAGAGAGTGTGCGCGGAGACACCGCC 4617
1218 eAlaGluProLeuProGluProLysArgLeuLeuAspAlaGln.Ar 1234
|||||
4618 TGGCTGAGCGCTGCTCCCTGAACCAAGAGAGAGCTGTCTGACGCTCAGG 4667
1234 gGlySerPheProProTrpValGlnGlnThrArgVal...ArgTrpProH 1250
|||||
4668 AGGCAGCTTCCCTTGGGTGCAAAACCGCGGTGTGACGCTGCGCCCTC 4717
1250 IsLeuGlyThrAlaTrpProGlnProHisProHisProHisProGlyCys 1266
|||||
4718 ACCTTGGAACGGCTGCGCCCGCCCGACCCCGCCCGCCCGCGCTGC 4767
1267 ArgSerGlnArgThrAlaSerSerGlyThrPro 1277
4768 AGATCAGAGAGCGCGAGTTCCGGAACACCG 4800

seq_name: gb_htg20:AL161903

seq_documentation_block:
LOCUS AL161903 150956 bp DNA HTG 29-SEP-2000
DEFINITION Homo sapiens chromosome 6 clone RP11-175A4, *** SEQUENCING IN
PROGRESS ***, 3 unordered pieces.
ACCESSION AL161903
VERSION AL161903.12 GI:10039702
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Tracy, A.
Direct Submission
Submitted (28-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerquest@sanger.ac.uk
On Sep 9, 2000 this sequence version replaced gi:9796876.
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: bA175A4
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 99% of reads
Chemistry: Dye-primer-amersham; 0% of reads
Chemistry: Dye-primer Big Dye; 0% of reads
Consensus quality: 145680 bases at least Q40
Consensus quality: 147560 bases at least Q30
Consensus quality: 148600 bases at least Q20
Insert size: 150756; sum-of-contigs
Insert size: 148326; 10.8% error; agarose-fp
Quality coverage: 6.43x in Q20 bases; sum-of-contigs Quality

```

```

coverage: 6.54x in Q20 bases; agarose-fp
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 54611: contig of 54611 bp in length
* 54612 54711: gap of 100 bp
* 54712 70293: contig of 15582 bp in length
* 70294 70393: gap of 100 bp
* 70394 150956: contig of 80563 bp in length.
FEATURES
source
1..150956
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="6"
/clone="RP11-175A4"
/clone.lib="RPC1-11.1"
misc_feature
1..54611
/note="assembly_fragment:01454"
vector_end:T7
vector_side:left
misc_feature
54712..70293
/note="assembly_fragment:02542"
fragment_chain:1"
70394..150956
/note="assembly_fragment:02844"
fragment_chain:1
vector_end:SP6
vector_side:right"
BASE COUNT 39575 a 37381 c 36411 g 37389 t 200 others
ORIGIN

alignment_scores:
Quality: 4066.00 Length: 3227
Ratio: 3.836 Gaps: 22
Percent Similarity: 32.848 Percent Identity: 31.887

alignment_block:
US-09-294-298-4 x AL161903/rev ..
Align seg 1/1 to reverse of: AL161903 from: 1 to: 150956

46 ArgValSerProGlyGlyLysGlnTyrSerMetGluAlaAa.ProAlaA 62
|||||
107790 CGCTTGAGCCGAGGAGCAGAGTTCGACGACGACGACGACCTG 107741
62 la.....ProPheArgProSerGlnGlyPheLeuSerArgLeuLys 76
:: ||| ||| |||:: |||:: |||
107740 TACTCCAGCCCTGGCGCAGACGAG..ACTGCTCAAAAAAAAAA 107694
77 SerSerIleLysArgThrLysSerGlnProLysLeuAspArgThrSerSe 93
::: ||| |||::: |||
107693 AAAAAAAAAAAGAGAGCTGGAGAGCC.....TGATCCCTTCC 107653
93 rPheArgGlnIleLeuProArgPheArgSerAlaAspHisAspArgAlaA 110
::: ||| |||
107652 TCTCCTGAACCTCTCTGCTGCC.....AGGGCCC 107624
110 rgLeuMetGlnSerPheLysGluSerHisSerHisGluSerLeuLeuSer 126
|||||
107623 GCGTGATGCAAAAGCTTTAAGGAGTACACTCTCATGAGTCTTGCTGAGT 107574
127 ProSerSerAlaAlaGluAlaLeuLeuLeuLeuLeuLeuLeuLeuLeu 143
|||||
107573 CCTAGCAGTGCAGCTGAGGCTTGGATGAGTCACTGGATGAGATTCAT 107524
143 eIleLysProValHisSerSerIleLeuGlnGlnPheCysPheGlu. 159

```

```

|||||
107523 TATCAAGCCAGTGCACAGCTCCATCTCGGGCCAGAGTCTCTGTTGAGG 107474
160 .Val..... 160
:::
107473 TACTGGGCTCGTGGGTGGGAGGCCAAAGACAGGGGTGATGGAAGG 107424
160 ..... 160
107423 TGGGGGCAGAGAGTCTAGAGAAAGTGGCAGAGTGGGACATCAAGGA 107374
160 ..... 160
107373 AACAGAACTTCTGGACTGGAGAAAGGTAGGCCAGGAGGAAGAGAA 107324
160 ..... 160
107323 GGTAGCAGAGTCTCTCCCTCTGTAGCCCTTTCCCTCAACTCCACACTC 107274
161 .....ThrThrSerSerGlyThrLysCysPheAlaCysArgSer 173
|||||
107273 CTTCCTAGGTAACAACCTCATCAGAACAAATGCTTTCCTGTCGGTCT 107224
174 AlaAlaGluArgAspLysTrpIleGluAsnLeuAlaAlaValLysPr 190
|||||
107223 GCGGCCGAAAGAGACAAATGGATTGAGATCTGCAGCGGCGCAATAAGCC 107174
190 oAsnLys..... 192
|||||
107173 CAACAAGTATTGGGAATAAAGGGGACACAACTGTGTAGGGCAAGGT 107124
192 ..... 192
107123 TGCACCAACACGGAAGCTGGGGCCCTAGGAGGAAGTGAGTTAAAGGA 107074
192 ..... 192
107073 GGAGAGGCTTGGGAAGSAGAGGATTGAGGTACAGTATCTGGACAAGC 107024
192 ..... 192
107023 AGGGGAGACCCCATTTATCTGAGTCCCCCATTTCTTTTCGTTCTGT 106974
192 ..... 192
106973 ACTGCTACCTGCCTTAGATCTCTTTCCCTGCCATAGATCAAGTACT 106924
192 ..... 192
106923 TACAGAGTTGATGGGCTCTGGAATTTCTTTAGTCTAGCTTCCCTGCAGG 106874
192 ..... 192
106873 CAGGAATGCTTCACTAATACCCGAGGACACATCAATACACTTAGCCA 106824
192 ..... 192
106823 AGTTTCTGACCTTGGTTTCTCTCTAAGACAGAGGGAAGTGGTGGCAAT 106774
192 ..... 192
106773 GGGTTGGATTAAATGATCTCTAAGTCCCTTTGGCACATAAATCTATGA 106724
192 ..... 192
106723 GTCTGTTCTTCCCAAACTTCATGCTTCCAGTTCTGTGACGATATCTCTG 106674
192 ..... 192
106673 TATCCAGTTTCTAGCAAGCTTACAGTCCCTAGTACACTCCCTCTGGGGG 106624
192 ..... 192

```

```

106623 AACTCTTGTGCTTGTATGTCCTTAAAGAACTCAACCCAGATGCTGACAT 106574
192 ..... 192
106573 GATCTGACCTATGCAGAGTACAAACACCATGTTCCCTTTCAACACACTGC 106524
192 ..... 192
106523 AATATCACTGAGCCTTACCATTAGCTTGTGCTCAATTTGTATCCCCCAG 106474
192 ..... 192
106473 GCCCTTACTGATTTCTTTTCTTTTCTTTTGTGACAGAGTCTCACTCTG 106424
192 ..... 192
106423 TCACCCAGGCTGGAGTGCAGTGGCGTGATCTTGGCTACTGCAACCTCCG 106374
192 ..... 192
106373 CCTCTGGGTTTAAAGCAATTCTCTGCTCAGCTCCGCCAGTAGCTGGGA 106324
192 ..... 192
106323 TTACAGGCATGCACACACACCCGACTAATTTTGTATTTTGTAGTAGAG 106274
192 ..... 192
106273 ACGAGTTTTCACCATGTGTGGCCAGGCTAGTCTCAAACTCCTGACCTCAGG 106224
192 ..... 192
106223 TTGCTGCTGCTGCCTCAGCTCCCAAGTGCTAGGATTACAGGCATGAGCC 106174
192 ..... 192
106173 ACCACATCTGCCCAAGCTTTACCCATTTCTATATGCAATCTTTTTCACA 106124
192 ..... 192
106123 ATTCTGATAGTCTCTGCAGGACTTCCAGTTCCCTTCAAAATCTACATT 106074
192 ..... 192
106073 AATATTTTGGCTTGTATTCCAGCTTTTGAATCTTTCCATACTGTTT 106024
192 ..... 192
106023 GTGTTGCTAGTGTATGTCTACGCTTCTACAGTAGCTTCCCTGGGGTGTCT 105974
192 ..... 192
105973 ATAACAACCTGCTACAACTGGTTCAGTAAACACAGAAATTTATTCT 105924
192 ..... 192
105923 CTCACAGTTCAGGAGGCTGGAAGTCCAAAGCAAGGTATCAGCAGGGCCA 105874
192 ..... 192
105873 CGCTCTCTTGTATGCTGTAGCAGGGAATCTCTCTTGGCTCTTCTAGCT 105824
192 ..... 192
105823 TCCGGAGGTTGCCAGCAGTCTTGGCATTTCTGGGCTTATAACTGCATCC 105774
192 ..... 192
105773 GTCTAATTTCTGCCTCCATCTTCATGTAGTGGCTTCTCTCTGTGTGTCT 105724
192 ..... 192
105723 CTGTATCCTGTATCTCTGTGTCTCCAAATCTCCCTCTCCATATAAGACA 105674

```

192 192
 105673 CCAGTTTTTATAGGTGGTTAAGGTCCTCACTCTAATTCAGTATGGCTC 105624
 192 192
 105623 TTTTTTTGAGATGGAATATCGCTCTTGTTCCTCCAGGCTTGGAGTGAAT 105574
 192 192
 105573 GGCATGAGTCAGCTCAGTCAACCTCCGCCTTCCAGGTTCAAGTGAATC 105524
 192 192
 105523 TCCTGCCTCAGCTCCTGAGTAGCTGGGATTAACAGGCACGTGCCACTATG 105474
 192 192
 105473 CCCAGCTAATTTGTTTGTATTTTATTATAGAGCGGGTTTCACTATGT 105424
 192 192
 105423 TGGCCAGGCTGGCTCGAAGCTCTGACCTCAGGCAATCCACCCAGCTCGC 105374
 192 192
 105373 CTCCCAAGTCTGGGATTACAGGAATGAGATACCATGGCTGGCTCTTC 105324
 192 192
 105323 TTAGTTGATTACATTTGCAGGATCCTGTTTCCAAATTAAGATCACATTC 105274
 192 192
 105273 ACAGTTTCAGGGTAGACATGAGTTTGGTGGGATCTATTCAACCCAGTA 105224
 192 192
 105223 CACCGCTCACCTTATGATTGCTGTCTGTATCCACCTTCTTACTGTCTC 105174
 192 192
 105173 TCCTCTTTTGGTGTCTCTGTCTCCTTCCCTTTAAACATCATGCCACC 105124
 193 205
 105123 CCACCATGCCAGGACACAGCCCGGGTAGACAATGTGCTAAAGCTGTG 105074
 205 pfileleGluAlaArgGluProLysLysArgTyrTyrCysGluL 222
 105073 GATCATAGGCGCGGAGCTGCCCGGAGCGGTACTACTGTGAGC 105024
 222 eucysLeuAspMetLeuTyrAlaArgThrThrSerLysProArgSer 238
 105023 TCTGCTGGATGACATGCTGTATGACGACGACACCTCCAAAGCCGCTCT 104974
 239 AlaSerGlyAspThrValPheTgGlyGluHisPheGluPheAsnLe 255
 104973 GCCTCTGGGACCGGCTCTCTGGGCGAGCACTTCGAGTTTAAACACT 104924
 255 uProAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysL 272
 104923 GCGGCTGTCGCTGCCCTGGGCTGCATCTGTACCGTGACTCAGACAAA 104874
 272 ysArgLysLysAspLysAlaGlyTyrValGlyLeuValThrProVal 288
 104873 ACCGCAAGACAGACAGCGAGGTATGTGCGCTGTGTACTGTGCCAGTG 104824
 289 AlaThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLe 305
 104823 GCCACCTGGCTGGCGCCACTTCACAGAGCAGTGTACCTGTAAACCT 104774

305 uProThrGlySerGlyGlySerGlyMetGlySerGlyGlyGlyG 322
 104773 GCCAAACAGGCAGTGGGGATCTGGGGCATGGGTTCGGAGGGGGAGGG 104724
 322 lysSerGlyGlySerGlyLysGlyLysGlyLysGlyLysGlyLys 338
 104723 GCTCGGGGGTGGCTCAGGGGCAAGGCAAGAGAGGTGGCCGGCTGTG 104674
 339 ArgLeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluLeu 355
 104673 CGGTGAAGACAGCTTACCAGACAATGAGCATCTTGCCTCAGGCTATA 104624
 355 rLysGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAla 372
 104623 TAAAGAGTTTGCAGAGTATGTCAACCACTTATCGATGCTGTGTGCG 104574
 372 alLeuGluProAlaLeuAsnValLysGlyLysGluGluValAlaSerAla 388
 104573 TCTTGGAGCCCGCTGAATGTCAAGGCAAGGAGGAGTTGCCAGTGCA 104524
 389 LeuValHisIleLeuGlnSerThrGlyLysAlaLys 400
 104523 CTAGTTCACTCTCGCAGAGTACAGCAAGGCAAGGTGAGTGTGTGCC 104474
 400 400
 104473 CTCAGGAAAGGTGACTTGGGAATGGGCACCTTGTGGGGTTAGTGAG 104424
 400 400
 104423 ACAGGCAANTTCAGGAGATTGGTGTGTCAGAGGCTGACACTTGGATT 104374
 401 AspPheLeuSerAspMetAlaMetSerGluValAspA 413
 104373 TCCTGGGCGCTCAGGACTTCTTTCACACATGGCCATGCTGAGGTAGACC 104324
 413 rGpMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThr 429
 104323 GGTTCATGGAACGGAGCACCTCATATTCGCGGAGAACACACGCTGCCACT 104274
 430 LysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuLys 446
 104273 AAAGCATAGAGAGTATATGAGACTGATTGGTCAGAAATACCTCAAGGA 104224
 446 pAla.Ile..... 448
 104223 TGCCATTGGTATGGCCACACTCAGGCCCTCTTCTTCCCAACCTGCCAG 104174
 448 448
 104173 ATGTCCACCCCAAGCCCAAGTCCACCTTCCACAGCTTGATACTTCCTA 104124
 448 448
 104123 ACCCAGAGTCTTAGGACTCCAGCCTCCAAACCTTGATTCGAAATTTCCC 104074
 448 448
 104073 CAACCTGGCCACCCCTTCCCTGCCCTTGGAAAGTGTGACCACACCCCTC 104024
 449 GlyGluPheIleArgAlaLeuTyrGluSerG 459
 104023 TTGTGCCCCCAACCCCAAGGAAATTCATCCGCTGCTGTATGAATCTGA 103974
 459 uGluAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaG 476
 103973 GGAACATGCGAGGTAGACCTTATCAAGTGCACAGCATCCAGTTTGGCAG 103924
 476 luHisGlnAlaAsnLeuArgMetCysGluLeuAlaLeuCysLysVal 492
 103923 AGCACCGGCGCAACCTCGGAATGTGTGAGTTGGCCCTGTGSCAAGGTG 103874
 493 ValAsn..... 494


```

Em:N71594 Em:AA095718 Em:AA377858 Em:H25235 Em:H52755
Em:D52710 Em:AA329903
/evidence=not_experimental
/product="dJ570F3.2 (POLYPEPTIDE FROM PATENTED CDNA
EMBL:E06811)"

alignment_scores:
  Quality: 4060.50      Length: 3227
  Ratio: 3.838         Gaps: 21
  Percent Similarity: 32.786 Percent Identity: 31.856

alignment_block:
US-09-294-298-4 x HSDJ570F3/rev ..

Align seq 1/1 to reverse of: HSDJ570F3 from: 1 to: 94770

46 ArgValSerProGlyGly.....LysGlnTyrSerMetGluAl 58
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10638 CGCTTGAGCCGAGGAGGAGAGITGCGACNAGCCGAGATCACCACCTG 10589

58 aAlaProAlaAlaProPheArgProSerGlnGlnPheLeuSerArgArgL 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10588 TACTCCAGCC.....TGGGCGACAGGCAAGACTGCTCAAAAAA 10545

75 eulysSerSerIleLysArgThrLysSerGlnProLysLeuAspArgThr 91
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10544 AAAAAA.....AAAGGAAGAGTGGGAAGCCC.....TGATCC 10507

92 SerSerPheArgGlnIleuProArgPheArgSerAlaAspHisAspArg 108
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10506 CTTCCTCTCTGACCTCTGCTGCC.....AG 10478

108 gAlaArgLeuMetGlnSerPheLysGluSerHisSerHisGluSerLeuL 125
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10477 GGCCCGGCTGATGCAAAGCTTTAAGAGAGTCACTCATGAGTCTCTGC 10428

125 euSerProSerSerAlaAlaGluAlaLeuGluLeuAsnLeuAspGluAsp 141
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10427 TGAGTCCTAGCAGTCACTGGAGCATTTGGAGCTCAACTGGATGAAGAT 10378

142 SerIleIleLysProValHisSerSerIleuGlyGlnGluPheCysPhe 158
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10377 TCCATTATCAACCAAGTGCACAGTCCATCTCTGGCCAGAGTTCGTGTT 10328

158 eGlu..Val..... 160
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10327 TGAGGTACTGGTCTGTGGTGGGCTGGGAGGCCAAAGACAGGGGTGATG 10278

160 ..... 160
10277 GAAGGTGGGGCGACAGAGGTCTAGAGAAAGTGCACAGGTGGGACATC 10228

160 ..... 160
10227 AAGAAACAGAACTCTGGGACTGGAGAAAGGGTAGGCCAGGGAGGAA 10178

160 ..... 160
10177 GACAAGGTAGCAGAGTCTCTCCCTCTGTAGCCCTTTCCCTCAACTCCA 10128

161 .....ThrThrSerSerGlyThrLysCysPheAlaCysA 172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10127 CACTCCTTTCTAGGTAACTTCTCCTGAGCAAAATGCTTTGCTGCTGC 10078

172 r9SerAlaAlaGluArgAspLysTrpIleGluAsnLeuGlnArgAlaVal 188
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10077 GGTCTCGGCCCAAGAGACAAATGGATTGAGATCTGCAGCGGGCAGTA 10028

189 LysProAsh..... 191
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
10027 AAGCCCAACAAGGTATTGGGAATAAAGGGGACACAACTGTGCAGGGCA 9978

191 ..... 191

```

9977 AAGGTTGCACCAACACGGGAAGCTGGGGCTAGGGAGGAAGTGAATTA 9928
191 191
9927 AAGGAGGAGAGGCTTGGGGAAGAGAGGATGAGGTACAGTGTATCTGGA 9878
191 191
9877 CAAGCAGGGGAGAGCCCAATTAATCTGAGTCCCAATTTCTTTTCGCTT 9828
191 191
9827 TCTGTACTGCTACCTCCCTTACGATCTCTTTCCCTGCCATAGAAGTCAAT 9778
191 191
9777 AGACTTACAGAGTTGATGGGGCTCTGGAATTTCTTTAGTCTAGCTTCCCT 9728
191 191
9727 GCAGCAGGAATGCTTCACTAATACCCCGCAGGACACATCAATACACTT 9678
191 191
9677 AGCCAAAGTTTCTGTACCTTGGTTTCTCTCTCTAAGACAGAGGAAGTGTG 9628
191 191
9627 GCAATGGGTGGATTAATGATCTCTAAGGTCCCTTTGGCACATAAATTC 9578
191 191
9577 TATGAGTCTGTCTTCCAAAACCTTCATGCTTCCAGTCTCTGTACGTATT 9528
191 191
9527 CTCGTATCCAGTTTCTAGCAAGCTTACAGTCTAGTACACACTCCTCTG 9478
191 191
9477 GGGGAAGTCTTGTCTGTATGCTCTTAAGAAGTCAACCAAGATGCT 9428
191 191
9427 GACATGATCTGACCTATGACAGGTACAAACACCAATGTTCCCTTCAACAC 9378
191 191
9377 ACTGCATATCACCTGAGCCTTACCATTAGCTGTGTGCTCAATTGTATCCC 9328
191 191
9327 CCCAGGCCCTTACTGATCTCTTTTTTTTTTTTGGACAGAGTCTCAC 9278
191 191
9277 TCTGTCAACCAAGCTGGATGAGTGGGTGATCTTGGCTCACTGCAACC 9228
191 191
9227 TCCGCTCTCGGTTTAAGCAATTCCTGTGCTTCCCTTCCAGTGTAGCT 9178
191 191
9177 GGGATTACAGGCATGCACCAACCAACCCGACCTTAATTTTGTATTTTAGT 9128
191 191
9127 AGAGACAGGTTTCAACATGTTGGCAGGCTAGTCTCAAACCTCCAGCT 9078
191 191

9077 CAGGTTGCTGCTGCTCAGCCTCCAAAAGTCTAGGATACAGCATG 9028
191 191
9027 AGCCACCACATCTGGCCAAGCTTTACCAATTTATATATGCAATTTCTTTT 8978
191 191
8977 CACAATTTCTGATAGTCTCTCAGGACTTTCCAGTTCCCTTCAAATCTA 8928
191 191
8927 CATTAATATTTTGGCTTGTATTTCCAGCTTTTGAATCTTTCCAATACT 8878
191 191
8877 GTTTGTGTCTCTAGTGTATGTCACGCTTCTACAGTACTTCTCTGGGC 8828
191 191
8827 TGCTATAACAAACTGCTACAAACTGGGTCAAGTCAAAACAACAGAAATTA 8778
191 191
8777 TTCTCTCACAGTTTCAGGAGCTGGAAGTCCAAAAGCAAGGTATACAGAG 8728
191 191
8727 GCCACGCTCTCTTTGATGCTGTAGCAGGGAATCCTTCCCTTGCCTCTTCT 8678
191 191
8677 AGCTTCCGAGGTTGCCAGCAGTCTTGCAATTCCTTGGCTTATAACTGC 8628
191 191
8627 ATCCGTCTAATTTCTGCTCCATCTCATGTAGTGGCTTCTCTGTGT 8578
191 191
8577 GTCCTGTATCCCTGTATCTGTGTCTTCCAATCTCCCTCTCCATATAA 8528
191 191
8527 GACACCAGTTTTTATAAGGTGGTTAAGGCTCCACTCTAATTCAGTATGG 8478
191 191
8477 CCTCTTTTTTTGAGATGGAATATCGCTCTTGTTCACAGCTTGGAGTG 8428
191 191
8427 CAATGGCATGAGTTCAGTCTACTGCAACCTCCGCTTCCAGGTTCAAAGTG 8378
191 191
8377 ATTCTCTGCTCAGCTCCTGTAGTGGGATTTACAGGCACGTGCCAC 8328
191 191
8327 TATGCCAGCTAATTTGTGTATTTTATTTAGACGGGTTTCACT 8278
191 191
8277 ATGTTGGCCAGGCTGGCTCGAACTCCTGACTCAGGCATTCACCCACCT 8228
191 191
8227 CGGCTCCCAAAGTGTGGGATTCAGGAATGAGATPACCATGGCTGCGCT 8178
191 191
8177 CTCTCTAGTTGATTACATTTTGCAGGGATCCTGTCTTCCAATAAGATCAC 8128

191 191
8127 ATTCACAGGTTTCAGGGTAGACATGAGTTTGTGGGATACCTATTCAACCC 8078
191 191
8077 AGTACACCGTCTCACCTTATGATTGCTGTCTGTATCCACCTTCTTACTG 8028
191 191
8027 TCTCTCTCTTTTGGTGTCTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 7978
192 191
192 Lys... AspAsnSerArgArgValAspAsnValLeuLysL 204
7977 CACCCACCATGCCAGGACCAACAGCCGCCGGGTAGACAAATGTCTTAAGC 7928
204 eutrpIlelleGluAlaArgGluLeuProProLysLysArgTyrTyrCys 220
7927 TGTGGATCATAGAGGCCCGGAGCTGCCCCCAAGAAGCGGTACTACTGT 7878
221 GluLeuCysLeuAspMetLeuTyrAlaArgThrThrSerLysProAr 237
7877 GAGCTCTGCTGATGACATGCTGTATGACGACACCTCCAAAGCCCG 7828
237 qSerAlaSerGlyAspThrValPheTrpGlyGluHisPheGluPheAsnA 254
7827 CTCTGCTCTGGGACACCGTCTCTGGGGCGGCGCTTCGAGTTTAAACA 7778
254 snLeuProAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAsp 270
7777 ACCTGCGGCTGCTGCGTCCGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 7728
271 LysLysArgLysLysAspLysAlaGlyTyrValGlyLeuValThrValPr 287
7727 AAAAAGCGCAAGAACAGCAGCGCTATGTCGCGCTGTGTGCTGTGCC 7678
287 oValAlaThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValT 304
7677 AGTGGCCACCTGGCTGGGCGCCACTTCACAGACAGTGGTACCCCTGTA 7628
304 hrLeuProThrGlySerGlySerGlySerGlyMetGlySerGlyGlyGly 320
7627 CCCTGCCAACACGACGTGGGGATCTGGGGGCATGGGTTCGGAGGGGGA 7578
321 GlyGlySerGlyGlySerGlyGlyLysGlyLysGlyGlyCysProAl 337
7577 GGGGGCTCGGGGGTGGCTCAGGGGCAAGGGCAAGGAGGTGGCCGGC 7528
337 aValArgLeuLysAlaArgTyrGlnThrMetSerIleLeuProMetGluL 354
7527 TGTGCGCTCAAGACAGCTTACCAGACATGAGCATCTTGCCCATGGAGC 7478
354 eutYrLysGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCys 370
7477 TATATAAGAGTTTGAGAGTATGTCACCAACCATATTCGGATGCTGTGT 7428
371 AlaValLeuGluProAlaLeuAsnValLysGlyLysGlyGluValAlaLase 387
7427 GCAGTCTTGGAGCCGCCCTGAATGTCAAAAGGCAAGGAGGAGGTGGCCAG 7378
387 rAlaLeuValHisIleLeuGlnSerThrGlyLysAla Lys 400
7377 TGCACCTAGTTTCATCTCTGCAGATACAGGCAAGGCCAAGGTGAGTGTG 7328
400 400
7327 TGCCCTCAGGAAAGGTGACTTTGGAAATGGGCACTTGCTTGGGGGTAGT 7278
400 400
7277 GAGGACAGGCAAAATTCACAGATGGGTTGTGCAGAGGCTGACACTGG 7228

401 AspPheLeuSerAspMetAlaMetSerGluVal 411
7227 ATTTTCCTGGGCTTCAGGACTTCCTTTTCAGACATGGCCATGTCTGAGGTA 7178
412 AspArgPheMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAl 428
7177 GACCGTTTCATGGAACGGGAGCACCTCATATTCCGCGAGAACACACGCTTGC 7128
428 aThrLysAlaIleGluGluTyrMetArgLeuIleGlyGlnLysTyrLeuL 445
7127 CACTAAAGCCCATAGAAGAGTATATGAGACTGATTGGTCAGAAATACCTCA 7078
445 ysAspAla.Ile 448
7077 AGGATGCCATTGGTATGGCCCACTCAGGCCCTCTCTCTCCAAACCTG 7028
448 448
7027 CCAGATGCCACCCACAGACCCCAAGTCCACCCTTCCACAGCTTGATACTT 6978
448 448
6977 CTAACCCAGAGTCTTAGGACTCCAGCTCCAAACCTGATTTCTGAAAT 6928
448 448
6927 TCCCAACCCCTGGCCACCCCTTCCCTGGCCCTTGGAAAGTGTGACACAC 6878
449 GlyGluPheIleArgAlaLeuTyrGluS 458
6877 CCTCTGTGCCCCACCCCGGAGGAAATTCATCCGTGCTCTGTATGAAT 6828
458 erGluGluAsnCysGluValAspProIleLysCysThrAlaSerSerLeu 474
6827 CTGAGGAAACTGCGAGTAGAGCTATCAAGTGCACAGCATCCAGTTG 6778
475 AlaGluHisGlnAlaAsnLeuArgMetCysCysGluLeuAlaLeuCysly 491
6777 GCAGAGCACCCAGGCCAACCTCGCAATGTGCTGTGAGTTGGCCCTGTGCAA 6728
491 sValValAsnSerHis 496
6727 GGTGTCAACTCCACCTGGTGAGACTGGGAACCTGGGCTGGGGGGCCAG 6678
496 496
6677 GGTGGGGGAATATGTGTTCATCTGTTCATCTGTTCCTCATCTCAAA 6628
496 496
6627 GAGGACTGAGCACCATTTATGGGCAAGCATTTGTCTTAGGCGCTATAGAG 6578
496 496
6577 CAACAGGTGAAGAGGCTGTGCTCCCTGCCCTCAGAGGGCTCCACCAGA 6528
496 496
6527 ATGGGCAAAATTAGAAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6478
496 496
6477 GTGTAAGTGTGAGTAAGGTGCCCTCCCTCTGTGTGACATAAGGTCA 6428
496 496
6427 GAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 6378
496 496
6377 AGGCTTTGAGACTGGTTAGGTTCTGATATATGACATAGTTGGCAAG 6328
496 496

6327 AAGACATTTTCAGAAAGGCTGTGTGAGAGGCACATGTGTGATGTGAAA 6278
 496 496
 6277 AGGCCAGGAGTTTTCAGGGGACATTAAGTAGTTAGTACGAATTACAT 6228
 496 496
 6227 CAGGTTTAGTGGAGCATGTGCCTCATATAATGGGAGTGGCGGAGAGATGT 6178
 496 496
 6177 CTGGCAGGAGATTAGCTTTAGAACTGGAAGGCTCAAGGAGTCTGAG 6128
 496 496
 6127 CTCATTGTGTAGGCTTGGGATGCCATTAAAGGTGTCAAGAAATATTGTGAT 6078
 496 496
 6077 TTAGAAGATTAAATCCATAGCTGGGCACCGTGGCTCACACCTGTAAATACC 6028
 496 496
 6027 AGCACTTTGGGAGGCTGAGCGGGCAGACCCTGAGGTCTAGGAGTTTGA 5978
 496 496
 5977 GACCAGGCTGACCAACATGGAGAAACCCCTGTCTCTACTAAAAATACAAAA 5928
 496 496
 5927 TTAGCCAGATGTGGTGGACATCCCTGTATAATCCAGCTATTGGGAGGCT 5878
 496 496
 5877 AAGCAGGAGAAATCACTTTGAATCTGGGAGTGGAGGTGTAGTAGAGCGA 5828
 496 496
 5827 GATCACAGGATTGCACCTTAGCCTGGGCAACAGAGTGAACCTCCATCTC 5778
 496 496
 5777 AAAAAAAAAAATTAATCCATAGTAGAGTGTATGTAGTAGAGTAGAGGG 5728
 496 496
 5727 ATGTATGTTGGGGATGACTGGAAATAGAGCGAGCTGGGGGCTCGGAGGCA 5678
 496 496
 5677 TGAAGTCAGATCCTGAATCAGAACAGTACAGAGTTAGGAAGAGCTA 5628
 496 496
 5627 CTGGAGAGCCCTCTCGGAAAGGATCAGTAGGAATTGTACGCTTATTGG 5578
 496 496
 5577 AAAGGAGGAGCATGTCTGTGGGAGTCAAGGATGACTCAAGAGGCCATG 5528
 496 496
 5527 AGGCTGGTGGTGGAGACCGGTGGCCTCATTCACAAACAGGAAGTACAG 5478
 496 496
 5477 GAGGAGAGCCAGTTGGAGGTGTGGGGGCGATGGTAGCTCTGCTTTTACA 5428
 496 496

5427 CAGCCAGGTTTAAGGTGTACGTGGGACATTTGAAGTGAAGTGGAGGTG 5378
 496 496
 5377 CGGGAGTACCTCTGAGCCATTCCAGGACTGGGGATCATGCCTGGGGC 5328
 496 496
 5327 ACCTCCATCCCATTTCCCTGGAATCCAGAGAGTTGGGGGTCCGAGCT 5278
 496 496
 5277 CCCTGTACCTCAAGTGACCTCCATCTCTCTCCCATCTCTGTCTCTCCT 5228
 496 496
 5227 GGTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5178
 496 496
 5177 CATCTCTCTCCAGGTGTCTCTCCCTCACCTTCTCTCTCTCTCTCTCTCT 5128
 496 496
 5127 TCTCTCCCTAATCTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5078
 496 496
 5077 GCCTCCCATCTTGTCTCTGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5028
 496 496
 5027 GTTTCACACCTCACCTCTCTACCAACCCCTCAGCATGTTCCCTGGAAG 4978
 496 496
 4977 CTGAGGTCTCTGGGCTCAGTCCCGTCTCTCTCTCTCTCTCTCTCTCTCTCT 4928
 497 497
 4927 CTCTGTCTCCCGACCTTCCCGGAGCTGTCTCCGAGGAGCTGAAG 4878
 505 luAlaPheAlaSerTrpArgLeuArgCysAlaGluArgGlyArgGluasp 521
 4877 AGGTGTTTGTCTGTGGGCTGTCTGTGGCAGAGCGAGCGGGAGGAC 4828
 522 ileAlaAspArgLeuIleSerAlaSerLeuPheLeuArgPheLeuCyspr 538
 4827 ATCGAGACAGGCTTATCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4778
 538 oAlaIleMetSerProSerLeuPheGlyLeuMetGlnGluTyrProAspg 555
 4777 AGCGATTATGTGCGCCAGTCTCTTTGGGCTTATGACAGAGTACCCAGATG 4728
 555 luGlnThrSerArgThrLeuIleAlaLysValIleGlnAsnLeu 571
 4727 AGCAGCTCAGGAACCTCACCTCATTTGCCAAGGTATCCAGAACCTG 4678
 572 AlaAsnPheSerLys 576
 4677 GCCAACTTTTCCAAAGTGGGAGGCTTCAGAGTGGGCGAGGCGAGGAGT 4628
 576 576
 4627 GGCAGGCGAGGAGTGGCAGGCTGGGGTTCGGAAGAGGCTCTCTCTGA 4578
 576 576
 4577 GTCCCCAGAGATCTGTAGATGGGAGGCTATGATACCTTGTGTGTGTG 4528
 576 576
 4527 TATGT 4478

577PheThrSerLysGluAspPheLeuGlyP 586
 4477 ATCTTCTGCAATCTTGTGCTAGGTTTACCTCAAGAGGACTTCTCTGGGCT 4428
 586 heMetAsnGluPheLeuGluLeuTrpGlySerMetGlnGlnPheLeu 602
 4427 TCATGAATGAGTTCTGAGACTGGATGGGTTCCATGACGAGTTTGG 4378
 603 TyrGluLeuSerAsnLeuAspThrLeuThrAsnSerSerPheGluG1 619
 4377 TATGAGATCTCCAAATCTGGACAGCTTCCACACTGCATGCCCTACTCTGGG 4328
 619 yTyrIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpG 636
 4327 TTACATCGACTTGGCGGAGAGCTCTCCACACTGCATGCCCTACTCTGGG 4278
 636 luValLeuProGlnLeuSerLys..Glu..... 644
 4277 AGGTGCTGCCCGCAGCTCAGCAAGGTGACGAGATCCCTCTTTGGCCTATC 4228
 644 644
 4227 CCCAGATGGCTCCAGAGTTTCTTGAGCCTTGAGAACTACCCTTTGAAGA 4178
 644 644
 4177 TTTTCTTCTCCCTTCTTCTCGAGGTGTCACCACTACTATCCCAATC 4128
 644 644
 4127 AGGCGCCCTCCACCTGCACCTCAGAGGCCCTCTTAGAGCTGGCACAGA 4078
 645AlaLeuLeuLysLeuGly 650
 4077 GCCCGCAGGTAAACAGCTCACCCTCCAGAGGCCCTCTGAAGCTGGGT 4028
 651 ProLeuProArgLeuLeuSerAspIleSerThrAlaLeuArgAsnProAs 667
 4027 CCAGTGCCTGGCTCTCTCAACAGCATCAGCACAGCTCTCAGGAACCCCAA 3978
 667 nIleGlnArgGlnProSerArgGlnSerGluArgAlaArgSerGlnProm 684
 3977 CATCAAGAGGAGCCAGAGCCAGAGTGGAGGCGCCCGCTCAGGCTG 3928
 684 etValLeuArgGlyProSerAlaGluMetGlnGlyTyrMetMetArgAsp 700
 3927 TGGTACTGCGGGGCCATCGCTGAGATGAGGGCTTACATGATGCGGGAC 3878
 701 LeuAsn..... 702
 3877 CTCACAGGTGAGCACCCTGGAGACGCCAGGCTGTGCCCTAGGAGCCCT 3828
 702 702
 3827 TCTCTATTTAGATATCTCTAGTGGGCCACATGATCTCTAGG 3778
 702 702
 3777 CTTGAAGAAGGAGGAGAAAGAACCAAGTTCTCAGGGGAGACATAAG 3728
 702 702
 3727 AGACAGGTACATGCTAGTGGTAGGCTGAGAGCCCTTTACAGCCTGAGGAG 3678
 702 702
 3677 TGAGAGATTTGGAGCTCTAGGAATAGGCTGAGGCTCCACCAACTCAGG 3628
 702 702
 3627 CTTAGTTGTAAGCCTAGAGCATCCCTGCTGCAAGCTCTGATTTGCTGCC 3578

702 702
 3577 CTCTGCTGCCCATCTAGTCCCGAGGCTGAGGTTTCAGCCAGCATGTCAAT 3528
 702 702
 3527 GTACGCCATGTGTCAAAATGTTCAAAATCTCAGTANTAGTACTAGTAATA 3478
 702 702
 3477 AGCACTTCCCGAGCCCGGACCAACACACAGACCTCCCATGATCC 3428
 702 702
 3427 AGCACTTAGACAGTGCACAGCAGAGCCTAGAGGGCTTCAGCCTGCTC 3378
 702 702
 3377 CAGAGTCCCGAGCCTCCATCTGATAGTGGTGCCCGTGTTCATTGCTG 3328
 702 702
 3327 CTCAATATTTGATGATGGCCCTGCTTTTCCCTCTGCCCGTGTCTTCTC 3278
 702 702
 3277 CATGACACATACCATCCACCATCCCGCTCTCTTTCATTGTTCCA 3228
 703S 703
 3227 CATCTCTCTCTCTCTGCTGTGCTGCCCTCTTTCATCTCTCTCCA 3178
 703 erSerIleAspLeuGlnSerPheMetAlaArgGlyLeuAsn..... 716
 3177 GCTCATCGACCTTCAGTCTTCATGGCTCAGGCCCTCAACAGGTGAGG 3128
 716 716
 3127 GCTCTCCCTCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3078
 716 716
 3077 CTCAGTGGCTTGGCCCTACTCTCTCTCTCTCTCTCTCTCTCTCTCT 3028
 716 716
 3027 TCTCTCCATATGTGCCAGCCCTGCCCATCCCTCTCTCTCTCTCTCTCT 2978
 716 716
 2977 CATCTCCCTCTCTAGGCTCACCCCTTCCCGGAGGGCCCTGTCTT 2928
 716 716
 2927 TCCCTTTACTCACCTGTCCCTCCCATCTCTCTCTCTCTCTCTCTCT 2878
 716 716
 2877 GCTGCCAGGCTAGCTCTAGCCCTTCCCTCTGGGTCCCATCTTTTCA 2828
 716 716
 2827 CAAGCCTGTGCCAGACACACAGAGGTTCAATTTGTAGGAGCCCTGACC 2778
 717Ser.SerMetAspMet 721
 2777 TTACCTTCTGCTGTGTGCCCTTCCCTCTCTCTCTCTCTCTCTCTCT 2728
 722 AlaArgLeuProSerProThrLysGluLysProProProProProG1 738
 2727 GCTGCCCTCCCTCTCCCAACAGGAGGAGGAGGAGGAGGAGGAGGAG 2678
 738 yGlyGlyLysAspLeuPheTyrValSerArgProProLeuAlaArgSerS 755


```

|||||
2677 TGGTGGTAAAGACCTGTTCTATGTAAGCGTCCACCCCTGGCCGCTTCCT 2628
755 erProAlaTyrCysThrSerSerSerAspIleThrGluProGluGlnLys 771
2627 CACCAGCATCTGACGAGCGCTCGGACATCACAGCGCCAGAGAGAG 2578
772 MetLeuSerValAsnLysSerValSerMetLeuAspLeuGlnGlyAspG1 788
2577 ATGCTGAGTGTCAACAAGAGTGTCTCCATGCTGAGCTTACAGGGTGATGG 2528
788 yProGlyClyArgLeuAsnSerSerValSerAsnLeuAlaAlaValG 805
2527 GCCTGGTGGCGCTCAACAGCAGCAGTGTTCGAACCTGGCGGCCGTAG 2478
805 lAspLeuLeuHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeu 821
2477 GGGACCTGCTGCATCAAGCAGCGCTCGCTGACAGCAGCTTGGGGCTA 2428
822 ArgProAlaProAlaGlyArgLeuSerGlnGlySerGlySerIleTh 838
2427 CGGCTGGCGCTGCCGAGCGCTCTCCAGGGGAGTGGCTCATCATCAC 2378
838 rAlaAlaGlyMetArgLeuSerGlnMetGlyValThrThrAspGlyValP 855
2377 GCGGCTGGCATGGCGCTCAGCCAGATGGGTGTCCACACAGCGGTGTCC 2328
855 roAlaGlnGlnLeuArgIleProLeuSerPheGlnAsnProLeuPheHis 871
2327 CTGCCCAGCAACTCGGAATCCCTCTCTCCAGAACCCCTCTCTTCCAC 2278
872 MetAlaAlaAspGlyProGlyProProAlaGlyHisGlySerSerG1 888
2277 ATGGCTGCTGATGGCGCAGTCTCCCGAGCGGCCATPGAGGGGGGGTGG 2228
888 yHisGlyProProSerSerHisHisHisHisHisHisHisHisHis 905
2227 CCATGGGCCACTTCTCTCCATCACCACCACCACCACCACCACCACCACC 2178
905 rgGlyGlyGluProProGlyAspThrPheAlaProPheHisGlyTyrSer 921
2177 GAGGTGGAGAGCCCTTGGGACACCTTTGCCCATCTCCATGGCTATAGC 2128
922 LysSerGluAspLeuSerThrGlyValProLysProProAlaAlaSerI1 938
2127 AAGAGTGAGGACCTCTCTCCGGGGTCCCAAGCCCTGCTGCTCCAT 2078
938 eLeuHisSerHisSerTyrSerAspGluPheGlyProSerGlyThrAspP 955
2077 CCTTCATAGCCACAGCTACAGTATGAGTTGGACCTCTGGCAGTACT 2028
955 heThrArgArgGlnLeuSerLeuGlnAspAsnLeuGlnHisMetLeuSer 971
2027 TCACCGCTGGCAGCTTCTACTCCAGGACACCTGCAGCAGCATGCTGCC 1978
972 ProProGlnIleThrIleGlyProGlnArgProAlaProSerGlyProG1 988
1977 CCTCCAGATCACCATGCTGCCAGAGCCAGCCCTCTCAGGGCCCTGG 1928
988 yGlyGlySerGlyGlySerGlyGlyGlyGlyGlyGlyGlnProProp 1005
1927 AGGTGGGAGCGGTGGGGCAGCGTGGGGTGGGGGGCCAGCGCCCTC 1878
1005 roLeuGlnArgGlyLysSerGlnGlnLeuThrValSerAlaAlaGlnLys 1021
1877 CATTCACAGGGGCAAGTCTCAGCAGTTGACAGTCAGCGCAGCCAGAAA 1828
1022 ProArgProSerSerGlyAsnLeuLeuGlnSerProGluProSerTyrG1 1038
1827 CCCCAGGCGCATCCAGCGGAATCTATTGCACTGCCAGACCCAGTATGG 1778
1038 yProAlaArgProArgGlnGlnSerLeuSerLysGluGlySerIleGlyG 1055
|||||

```

```

1177 CCCGCCCGTCCACGGCAACAGAGCCTCAGCAAGAGGAGGCGAGCATTTGGG 1728
1055 lySerGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 1071
1727 GCAGCGGGGACAGCGGTGGCGAGGGGGTGGGGGCTGAAGCCCTCCATC 1678
1072 ThrLys 1073
1677 ACCAAGCAGGTAGGTCAAGCAGCAGGAGGAGGAGGCGGCTGGGTCAACAACAGG 1628
1073 1627 GAGGGAAGAGAGATGGGGGTGGGGTTGAAACAGAGTCTGTGGCCGTA 1578
1073 1577 AGTTACAATCTTCTTGCCCTCTTGGCCATTAAACAAATGTGTATAGAG 1528
1073 1527 GGCCTGCTATGCCATGTGCTATGCCAGGCACCTAAGGATACGGCAGCTGA 1478
1073 1477 ACCCTCTTAGCAGCTCTCATTTCCAGAGGGGATTAATCCATAAGTAGAATGG 1428
1073 1427 GGGTGACTGGAATAGGCCAGCTGGGGGCTTGGAGGCGATGAAGATCAGAT 1378
1073 1377 CCTGAATCAGAACACAGGAGAAATTTCTTGGCATTTCAACTCATCTCT 1328
1073 1327 GGCATTTCAGAGGTGATAAGAGCTTACAGCAGGCGCAATAGGACTGGGACTGT 1278
1073 1277 CGTCTCTTTGGCTGTCTGCTTGGCCCTCTAATGTACCTGCTTTTCCC 1228
1074 1227 ACCTTTCTTTCTGCTGTCGCCCTCACTGCGCTTGTCCAGCATTTCTCA 1178
1077 nThrProSerThrLeuAsnProThrMetProAlaSerGluArgThrValA 1094
1177 GACACCATCCACATTGAACCCCAATGCCAGCTCTGAGCGGACAGTGG 1128
1094 laTtpValSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHis 1110
1127 CTTGGGCTCCACATGCCCTCACTCTCGGCTGACATCGAGAGTGGCCAC 1078
1111 lIleGluArgGluGluTyrLysLeuLysGluTyrSerLysSerMetAspG1 1127
1077 ATCGAGCGGGAAGAGTACAAGCTCAAGGAGTACTCAAAATCGATGATGA 1028
1127 uSerArgLeuAspArgValLys 1134
1027 GAGCGCGCTGGATAGGTACGA 1006

```

seq_name: gb_htg20:AL161906

```

seq_documentation_block:
LOCUS AL161906 169323 bp DNA HTG 18-AUG-2000
DEFINITION Homo sapiens chromosome 6 c-clone RP11-567N9, *** SEQUENCING IN
PROGRESS ***, 24 unordered pieces.
ACCESSION AL161906
VERSION AL161906.5 GI:9863619
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE human.
ORGANISM Homo sapiens

```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 169323)

REFERENCE AUTHORS TITLE JOURNAL

Submitted (17-AUG-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Aug 21, 2000 this sequence version replaced gi:9212890.

COMMENT

----- Genome Center

Center: Sanger Centre

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

----- Project Information

Center project name: BA567N9

----- Summary Statistics

Assembly program: XGAP4; version 4.5

Sequencing vector: pGAP4; 108752; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Consensus quality: 155424 bases at least Q40

Consensus quality: 160003 bases at least Q30

Consensus quality: 163252 bases at least Q20

Insert size: 167023; sum-of-contigs

Insert size: 186175; 5.7% error; agarose-fp

Quality coverage: 3.59x in Q20 bases; sum-of-contigs Quality

coverage: 3.22x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 24 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 5456: contig of 5456 bp in length
* 5457 5556: gap of 100 bp
* 5557 8767: contig of 3211 bp in length
* 8768 8867: gap of 100 bp
* 8868 17168: contig of 8301 bp in length
* 17169 17268: gap of 100 bp
* 17269 26591: contig of 9323 bp in length
* 26592 26691: gap of 100 bp
* 26692 31188: contig of 4497 bp in length
* 31189 37579: contig of 6291 bp in length
* 37580 37679: gap of 100 bp
* 37680 45350: contig of 7671 bp in length
* 45351 45450: gap of 100 bp
* 45451 48354: contig of 2904 bp in length
* 48355 48454: gap of 100 bp
* 48455 63401: contig of 14947 bp in length
* 63402 63501: gap of 100 bp
* 63502 67724: contig of 4223 bp in length
* 67725 67824: gap of 100 bp
* 67825 70550: contig of 2726 bp in length
* 70551 70650: gap of 100 bp
* 70651 80664: contig of 10014 bp in length
* 80665 80764: gap of 100 bp
* 80765 83307: contig of 2543 bp in length
* 83308 83407: gap of 100 bp
* 83408 88383: contig of 4976 bp in length
* 88384 88483: gap of 100 bp
* 88484 93697: contig of 5214 bp in length
* 93698 93797: gap of 100 bp
* 93798 99354: contig of 5557 bp in length
* 99355 99454: gap of 100 bp
* 99455 111055: contig of 11601 bp in length
* 111056 111155: gap of 100 bp
* 111156 118645: contig of 7490 bp in length
* 118646 118745: gap of 100 bp
* 118746 122421: contig of 3676 bp in length

* 122422 122521: gap of 100 bp
* 122522 132468: contig of 9947 bp in length
* 132469 132568: gap of 100 bp
* 132569 137959: contig of 5391 bp in length
* 137960 138059: gap of 100 bp
* 138060 151691: contig of 13632 bp in length
* 151692 151791: gap of 100 bp
* 151792 158253: contig of 6462 bp in length
* 158254 158353: gap of 100 bp
* 158354 169323: contig of 10970 bp in length.

FEATURES

source

1. 169323

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="6"

/clone="RP11-567N9"

/clone_lib="RPC1-11.2"

1. 5456

/note="assembly_fragment:00884"

fragment_chain:1"

5557..8767

/note="assembly_fragment:00832"

fragment_chain:1"

8868..17168

/note="assembly_fragment:01060"

fragment_chain:1"

17269..26591

/note="assembly_fragment:01172"

fragment_chain:1"

26692..31188

/note="assembly_fragment:00279"

fragment_chain:1"

31289..37579

/note="assembly_fragment:01205"

fragment_chain:1"

37680..45350

/note="assembly_fragment:01334"

fragment_chain:2"

45451..48354

/note="assembly_fragment:01263"

fragment_chain:2"

48455..63401

/note="assembly_fragment:00990"

fragment_chain:2"

63502..67724

/note="assembly_fragment:00045"

fragment_chain:2"

67825..70550

/note="assembly_fragment:00965"

fragment_chain:2"

70651..80664

/note="assembly_fragment:00637"

fragment_chain:3"

80765..83307

/note="assembly_fragment:01291"

fragment_chain:3"

83408..88383

/note="assembly_fragment:01317"

fragment_chain:3"

88484..93697

/note="assembly_fragment:01171"

fragment_chain:4"

93798..99354

/note="assembly_fragment:00288"

fragment_chain:4"

99455..111055

/note="assembly_fragment:00286"

fragment_chain:4"

111156..118645

/note="assembly_fragment:00303"

fragment_chain:5"

118746..122421

/note="assembly_fragment:01184"

```

misc_feature      fragment_chain:5"
122522..132468
/note="assembly_fragment:00501
fragment_chain:6"
misc_feature      132569..137959
/note="assembly_fragment:01244
fragment_chain:6"
misc_feature      138060..151691
/note="assembly_fragment:00032"
misc_feature      151792..158253
/note="assembly_fragment:00779"
misc_feature      158354..169323
/note="assembly_fragment:01270"
BASE COUNT 41674 a 41332 c 41269 g 42673 t 2375 others
ORIGIN

alignment_scores:
  Quality: 3987.00      Length: 2457
  Ratio: 4.237          Gaps: 12
Percent Similarity: 38.299 Percent Identity: 37.810

alignment_block:
US-09-294-298-4 x AL161906 ..
Align seg 1/1 to: AL161906 from: 1 to: 169323

190 ProAsnLysAspAsnSerArgValAspAsnValLeuLysLeuTrpI1 206
|||||
77990 CCATGCCAGGACACACGCCCGGGTAGACAATGTCTAAAGCTGGAT 78039

206 eileGluAlaArgGluLeuProProLysLysArgTyrTyrCysGluLeuC 223
|||||
78040 CATAGAGCCCGGAGCTGCCCGCCCAAGAGCGGTACTGTGAGCTCT 78089

223 ysLeuAspAspMetLeuTyrAlaArgThrThrSerLysProArgSerAla 239
|||||
78090 GCCTGGATGACATGCTGTATGACGACGCCACCCTCCAAAGCCCGCTCTGCC 78139

240 SerGlyAspThrValPheTrpGlyGluHisPheGluPheAsnAsnLeuPr 256
|||||
78140 TCTGGGACACCGCTCTCTGGGGCGAGCACCTTCGAGTTTAAACCTGCC 78189

256 oAlaValArgAlaLeuArgLeuHisLeuTyrArgAspSerAspLysLys 273
|||||
78190 GCCTGTCCGTGCCCTGCGGCTGCATCTGTACCGTGACTCAGACAAAAGC 78239

273 rGlyLysAspLysAlaGlyTyrValGlyLeuValThrValProValAla 289
|||||
78240 GCAAGAGGACAGGACAGGCTAIGTCGGCTGTGACTGTGCCAGTGCC 78289

290 ThrLeuAlaGlyArgHisPheThrGluGlnTrpTyrProValThrLeuPr 306
|||||
78290 ACCCTGGCTGGGCGCCACTTCACAGAGCAGTGTGTACCTGTAAACCTGCC 78339

306 oThrGlySerGlyGlySerGlyGlyMetGlySerGlyGlyGlyGlyGlyS 323
|||||
78340 AACAGGACAGTGGGGATCTCGGGGCGATGGTTCCGGAGGGGGGGGCT 78389

323 erGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 339
|||||
78390 CGGGGGTGGCTCAGGGGGGCAAGGGCAAGAGGAGGTGGCCGGCTGTGGCG 78439

340 LeuLysAlaArgTyrGlnThrMetSerIleLeuPrometGluLeuTyrLy 356
|||||
78440 CTGAAGCAGCGTTACACAGCAATGAGCATCTTGGCCCATGGAGCTATATA 78489

356 sGluPheAlaGluTyrValThrAsnHisTyrArgMetLeuCysAlaValL 373
|||||
78490 AGAGTTTGCAGAGTATGTCACCAACCATTTATCGGATGCTGTGTCAGGFT 78539

373 euGluProAlaLeuAsnValLysGlyGlyGlyGlyGlyGlyGlyGlyGly 389
|||||

```

```

78540 TGGAGCCCGCCCTGAATGTCAAAAGCAAGAGGAGGTTCACAGTGCAC 78589
390 ValHisIleLeuGlnSerThrGlyLysAla Lys..... 400
|||||
78590 GTTCACATCTCGAGAGTACAGCAAGGCCAAGGTGAGTGTGTGCGCTC 78639
400 ..... 400
78640 AGGAAAGGTGACTTGGGAATGGCAGCTTGTCTTGGGGGCTTAGTGAGGACA 78689
400 ..... 400
78690 GGSCAAATTCACGAGATTGGGTGTTCACAGAGCTCACACTTGGATTTC 78739
401 .....AspPheLeuSerAspMetAlaMetSerGluValAspArgP 414
|||||
78740 TGGGCTTCAGGACTTCTCTTCAGACATGGCCATGTCTGAGGTAGACGGT 78789
414 heMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThrLys 430
|||||
78790 TCATGGAACGGGAGCACCTCATATTCCGCGAGACACACGCTTGGCCATAA 78839
431 AlaIleGluLysTyrMetArgLeuIleGlyGlnLysTyrLeuLysAspAl 447
|||||
78840 GCCATAGAGAGTATATGAGACTGTGTGTCAGAAATACCTCAAGGATGC 78889
447 a.lle..... 448
|||||
78890 CATTGGTATGGCCACACACTCAGGCCCTCTTCTCCCAAACTGCCAGATG 78939
448 ..... 448
78940 TCCACCCACAGACCCCAAGTCCACCTTCCACAGCTTGATCTCTTAACC 78989
448 ..... 448
78990 CAGAGTCTTAGAGCTCCAGCTCCACACCTGATTCTGAAATTTCCCCAA 79039
448 ..... 448
79040 CCCTGGCCACCCCTTCCCTGGCTGGAAAGTGTGACCACACCCCTCTTG 79089
449 .....GlyGluPheIleArgAlaLeuTyrGluSerGluG1 460
|||||
79090 TGCCCCCACCACCCAGGAGAAATTCATCCGTGCTGTATGAATCTGAGGA 79139
460 uAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaGluH 477
|||||
79140 AAATGCGAGTAGACCTTATCAAGTGCACAGCATCCAGTTTGGCAGAGC 79189
477 isGlnAlaAsnLeuArgMetCysGlyLeuAlaLeuCysLysValVal 493
|||||
79190 ACCAGGCAACCTGCGAAATGTGCTGTGAGTTGGCCCTGTGCAAGGTGTC 79239
494 AsnSerHis..... 496
|||||
79240 AACTCCCACTGGTGAGACTTGGGAAACGCTGGGCTGGGGGCCAGGGTGGG 79289
496 ..... 496
79290 GGAATTATGTCTCATCTGTTCATCTATCTGTCCATCTCAAGAGGACT 79339
496 ..... 496
79340 GAGCACCATTTTGGGCAAGCATTTGTTCTAGGCGCTATAGAGCAACAG 79389
496 ..... 496
79390 GTGAAAGAGGCTGGTCCCTGCCCTCAGAGGGCTCCACAGATGGGGA 79439
496 ..... 496
79440 CAAATTAGAAAGAAAAAAGGCAAGCCACAGAGCCATAATGTTGTGTAAG 79489

```

```
496 ..... 496
79490 TGCTGAGTAAGGTTCCCCCAACCTCTGTGTGACATAAGGTCAGAGAGAA 79539
496 ..... 496
79540 GGCAGAGCTTTGAGATAAGTGGGGAAGAGGTGCCCCCTTTGGTAGGCTTT 79589
496 ..... 496
79590 GAAGACTGGTTTAGGTTCTGATATATGGACATAGTTGGCAAGAAAGACAT 79639
496 ..... 496
79640 TTCAGAAGAAGGCTGTGAGAAAGGCACATGTGTGATGGTGAAGGCCCA 79689
496 ..... 496
79690 GGAGTTTTTCAGGGGACATTAAAGTAGGTTAGTAGCAATTACATCAGGTTT 79739
496 ..... 496
79740 AGTGGAGATCTGCCATATAATGGGAGTGGCGGAGAGATGCTCGGCCA 79789
496 ..... 496
79790 GGAAGATTAGCTTTAGAAACTGGAAGGCTCAAGGAGTCTGAGGTCAATTG 79839
496 ..... 496
79840 TTAGGCTTGGGATGCCATTAAAGGTGTCAGAAATATTGTGATTAGAAG 79889
496 ..... 496
79890 ATTAATCATAGCTGGGCAGGTGGCTCACACCTGTAATACCACACTT 79939
496 ..... 496
79940 TGGAGGCTGAGGGGCGAGACCACCTGAGGTGAGGTTTGAGACCAGG 79989
496 ..... 496
79990 CTGACCAACATGGAGAAACCTCTCTACTATAAATAAATAATAGCCA 80039
496 ..... 496
80040 GATGTGGTGACATGCGCTGTATATCCAGCTATTCGGAGGCTAAGGCAG 80089
496 ..... 496
80090 GAGAAATCATTGAATCTGGGAGGTGGAGGTTGTAGTGAGCCGAGATCACA 80139
496 ..... 496
80140 CCATTGCACTAGCTGGGCAACAAGAGTGAACCTCCGTCFAAAAAA 80189
496 ..... 496
80190 AAAAAATTAATCAATAGTAGAGTGTATGTAGTAGAGTAGAGGATGTATG 80239
496 ..... 496
80240 TTGGGGATGACTGAATAGAGCCAGCTGGGGCTCGGAGGATGAAGT 80289
496 ..... 496
80290 CAGATCCTGTAATCAGAACAGTACAGAAAGTTAGGAAGGAGCTACTGGAAG 80339
496 ..... 496
80340 GACCCCTCTCGGAAAGGATCAGTAGGAATGTGACGTTATTGGAAGGAG 80389
```

```
496 ..... 496
80390 GGAGCATCTCTGTGGGAGTACGGATGACTCAAGAGGCCATGAGGCTGG 80439
496 ..... 496
80440 TGGTTGGAGAGCCGCTGCCTCATTTGACAAACAGGAAAGTCAGGAGGAGG 80489
496 ..... 496
80490 AGCCAGTTGAGGTGTGGGGCGATGCTGAGCTCTGCTTTACACCAGCG 80539
496 ..... 496
80540 AGTTTAAGGTCAGTGGGACATTGAAGTGGAACTGGGAGGTGCGGGGAG 80589
496 ..... 496
80590 TGAGCTCTGAGCCATCCAGGGACTGGGATCATGCCCTGGGCACTCCA 80639
496 ..... 496
80640 TCCCATTTCCCTGGATCCAGAGANNNNNNNNNNNNNNNNNNNNNNNNNN 80689
496 ..... 496
80690 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 80739
496 ..... 496
80740 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 80789
496 ..... 496
80790 GGTCCGAGCTCCCTGTACCTCAAGTGACCCCTCCATCTCTCTCCATCT 80839
496 ..... 496
80840 GTCTCTCCCTGGTGTCTGTTTTTCTCTCTCTCTCTCTCTCTCTCTCT 80889
496 ..... 496
80890 CACACCCCTCCATCTCTCTCCAGGTCTCTCCCTCACCCTCTCTCTCTCC 80939
496 ..... 496
80940 CCTCCATTTCTCTCTCCCTAATCTCTCTCTCTCTCTCTCTCTCTCTCT 80989
496 ..... 496
80990 TCTTCAAGCAGCTCCCATCTTGTCTCTGGGTCCCTCTCTCTCTCTCTCT 81039
496 ..... 496
81040 TCTACCCCTGTTTCCACACCCCTCACCTCTCTACACCCCTCAGCATGT 81089
496 ..... 496
81090 TCCCTGGAAGCTGAGGGTCTCTGGGGCTCAGTCCCGTCTCTCTCTCTCT 81139
497 ..... 501
81140 CTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 81189
502 GluLeuLysGluValPheAlaSerTrpArgLeuArgCysAlaGluArgG 518
|||||
81190 GAGCTGAAGAGGAGGTTTGTCTTATGGCGCTGCGTGGCAGAGCGAGG 81239
|||||
518 YArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuArgP 535
|||||
81240 CCGGAGGAGACATCGCAGACAGGCTTATCAGCGCTCACCTCTCTCTCTCT 81289
|||||
535 heLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGlnGlu 551
```

```
|||||
81290 TCCTCTGCCAGGATATGTGCCCGAGTCTTTGGGCTTATGCAGAG 81339
552 TyrProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysValI 568
|||||
81340 TACCAGATGAGCAGACCTCACGAACCTCACCTCATTCGCAAGGTCA 81389
568 eGlnAsnLeuAlaAsnPheSerIys.
|||||
81390 CCAGAACCTGCCAACTTTTCCAAGTGAGGGAAGCTTCAGAGTGGGCGAG 81439
576
81440 GGCAGGAGTGGCAGGCGAGGTGGCAGGCTGGGGTGGCAAGAAG 81489
576
81490 GGTCTCTGAGTCCCGAGAGATCTGAGATGGGAGGCTATGATACCTTG 81539
576
81540 TGTGTGTGTATGTGTGTATGTGTGTGTGTGTGTGTGTGTGTGTGT 81589
577
81590 TGTGACCTTTATCTTCTGCATTTCTGGCTAGGTTTACCTCAAGGAGGAC 81639
583 PheLeuGlyPheMetAsnGluPheLeuGluLeuGluTrpGlySerMetG 599
81640 TTCTGGGCTTCATGATGATGATGATGATGATGATGATGATGATGAT 81689
599 nGlnPheLeuTyrgLileSerAsnLeuAspThrLeuThrAsnSerSerS 616
81690 GCAGTTTTTGTATGAGATCTCCAACTGGACACGCTTAACCAACAGCAGTA 81739
616 erPheGluGlyTyrgLileAspLeuGlyArgGluLeuSerThrLeuHisAla 632
81740 GCTTTGAGGGTTACATCGACTTGGGCCGAGAGCTCTCCACACTGCATGCC 81789
633 LeuLeuTrpGluValLeuProGlnLeuSerLys..Glu.
|||||
81790 CTACTCTGGAGGTTGCTGCCCGCAGCTCAAGAGGTGAGAGTCCCTCT 81839
644
81840 TTGCCCTATCCCGATGGCTCCAGAGTTCTCTGGAGCTGAGAACTAC 81889
644
81890 CTTTGAAGATTTTTCCTCCCTTGTTCGAGGTGTCACCACTACT 81939
644
81940 ATCCCACTCAGGCCCTCCACCTGCACCTCAGAGGCCCTTTAGAGC 81989
645
81990 TGGGACTGAGCCCCAGGTAAACAGCTCACCCTCAGAGGCCCTTCCAG 82039
647
82040 GAACTGGGTCCACTGCGCGGCTCTCAACGACATCAGCAGCTCTGA 82089
664 rAsnProAsnIleGlnArgGlnProSerArgGlnSerGluArgAlaArg 680
82090 GGAACCCCAATCCAAAGGAGCCCAAGCGCCAGAGTGGGCGCCCGG 82139
681 SerGlnProMetValLeuArgGlyProSerAlaGluMetGlnGlyTrpMe 697
82140 CCTCAGCCTGTGGTACTGCGGGGCCATCGGCTGAGATCAGGGCTACAT 82189
697 tMetArgAspLeuAsn.
|||||
```


84940 TGAAGTACGGCCAGCTGGGGCTTGGAGGATGAGATCAGATCCTGTAAT 84989
 1073
 84990 CAGAACACAGAGAAATCTCTGGCATTCAACTACATCCTCTGGCATTC 85039
 1073
 85040 AGAGTGATGATAAGGCTACAGCAGGCAATAGGACTGGGACTGCTGCTCC 85089
 1073
 85090 TTGGCTGCTGCTGTTGGCTCTAAATGTACCTGCTTTTCCACCTTTT 85139
 1074
 85140 CTTTCTCTGTTGCGCCTCACTGCTGCTTGTCCAGCATTCCTCAGACCA 85189
 1080 SerThrLeuAsnProThrMetProAlaSerGluArgThrValAlaTrpVa 1096
 85190 TCCACATTGAACCCACACATGCCAGCTCTGACGGGACAGTGGCTGGGT 85239
 1096 lSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHisIleGluA 1113
 85240 CTTCAACATGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 85289
 1113 rGluGluTyrLysLeuLysGluTyrSerLysSerMetAspGluSerArg 1129
 85290 GGGAGAGTACACAGCTCAAGGAGTACTCAAAATCGATGATGAGACCGG 85339
 1130 LeuAspArgValLys 1134
 85340 CTGGATAGGGTACCA 85354

seq_name: gb_pr3:AF047711
 seq_documentation_block: 4368 bp mRNA
 LOCUS AF047711
 DEFINITION Homo sapiens nGAP mRNA, complete cds.
 ACCESSION AF047711
 VERSION AF047711.1 GI:4105588
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Noto,S., Maeda,T., Hattori,S., Inazawa,J., Inamura,M., Asaka,M. and
 Hatakeyama,M.
 A novel human RSCAP-like gene that maps within the prostate cancer
 susceptibility locus at chromosome 1q25
 FEBS Lett. 441 (1), 127-131 (1998)
 99093006
 2 (bases 1 to 4368)
 Noto,S., Maeda,T., Hattori,S. and Hatakeyama,M.
 Direct Submission
 Submitted (11-PEB-1998) Department of Viral Oncology, The Cancer
 Institute, Japanese Foundation for Cancer Research, 1-37-1
 Kami-Ikebukuro, Toshima-ku, Tokyo, Japan
 Location/Qualifiers
 1..4368
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1q24"
 126..3545
 /note="ras GTPase activating protein"
 /codon_start=1
 /product="nGAP"
 /protein_id="AA040481.1"
 /db_xref="GI:4105589"
 /translation="MQTPEVPAERSPRRRISGPTSTSEKPNMSMDTANTSPFKVPGFFS"

alignment_scores:
 Quality: 2603.00 Length: 1324
 Ratio: 2.902 Gaps: 42
 Percent Similarity: 67.749 Percent Identity: 45.317
 alignment_block:
 US-09-294-298-4 x AF047711 ..

Align seg 1/1 to: AF047711 from: 1 to: 4368

48 SerProGlyGlyLysGlnTyrSerMetGluAlaAlaProAlaAlaProPh 64
 186 AGTACATCAGAGAAACCCCACTCCATGGACACTGCAATACCTCACCCTT 235
 64 eArgProSerGlnGlyPheLeuSerArgArgLeuLysSerSerIleLysA 81
 236 CAAA...GTACCAGATCTTCAGCAAGCGCTGAAAGGCTCCATCACA 282
 81 rGThrLysSerGlnProLysLeuAspArgThrSerPheArgGlnIle 97
 283 GGACCAAAAGCCAGTCAAGAGCTTGACAAACACAGAGCTTCGG..... 326
 98 LeuProArgPheArgSerAlaAspHisAspArgAlaArgLeuMetGlnSe 114
 327 CTTCCATCCCTTCGACGATACAGAT...GACAGGCTCTCGTGGGCTGCTAA 373
 114 rPheLysGluSerHisSerHisGluSerLeuLeuSerProSerSerAlaA 131
 374 ACTAAAAGAGTACAGTTCCTCCCATGAATCTTCTGAGCCCATGAGCAG 423
 131 laGluAlaLeuGluLeuAsnLeuAspGluAspSerIleIleLysProVal 147
 424 TGAATGTCTGATCTTGTGTAGAGGCAACCTGTATCAGTGAACCACTT 473
 148 HisSerSerIleLeuGlyGlnGluPheCysPheGluValThrThrSerSe 164
 474 CATAGTACATCTTGGACAGACTCTCTGTTGAGGTTACTACTTAAG 523
 164 rGlyThrLysCysPheAlaCysArgSerAlaAlaGluArgAspLysTrpI 181
 524 TGGAGTAAATGCTTCAGCTTAATCTGCTTCTGAGAGAGACCAATGGA 573
 181 leGluAsnLeuGlnArgAlaValLysProAsnLysAspAsnSerArgArg 197
 574 TGGAAACCTTCGAGGAGGAGTTCACCTAATAAGGACAAATTCAGGCGA 623
 198 ValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGluLeuProPr 214
 624 GCTGAAATGTCTCTGCTTTATGATCATTTGAAGCAAGCAAGACCTTCC 673

84940 TGAAGTACGGCCAGCTGGGGCTTGGAGGATGAGATCAGATCCTGTAAT 84989
 1073
 84990 CAGAACACAGAGAAATCTCTGGCATTCAACTACATCCTCTGGCATTC 85039
 1073
 85040 AGAGTGATGATAAGGCTACAGCAGGCAATAGGACTGGGACTGCTGCTCC 85089
 1073
 85090 TTGGCTGCTGCTGTTGGCTCTAAATGTACCTGCTTTTCCACCTTTT 85139
 1074
 85140 CTTTCTCTGTTGCGCCTCACTGCTGCTTGTCCAGCATTCCTCAGACCA 85189
 1080 SerThrLeuAsnProThrMetProAlaSerGluArgThrValAlaTrpVa 1096
 85190 TCCACATTGAACCCACACATGCCAGCTCTGACGGGACAGTGGCTGGGT 85239
 1096 lSerAsnMetProHisLeuSerAlaAspIleGluSerAlaHisIleGluA 1113
 85240 CTTCAACATGCTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 85289
 1113 rGluGluTyrLysLeuLysGluTyrSerLysSerMetAspGluSerArg 1129
 85290 GGGAGAGTACACAGCTCAAGGAGTACTCAAAATCGATGATGAGACCGG 85339
 1130 LeuAspArgValLys 1134
 85340 CTGGATAGGGTACCA 85354
 seq_name: gb_pr3:AF047711
 seq_documentation_block: 4368 bp mRNA
 LOCUS AF047711
 DEFINITION Homo sapiens nGAP mRNA, complete cds.
 ACCESSION AF047711
 VERSION AF047711.1 GI:4105588
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Noto,S., Maeda,T., Hattori,S., Inazawa,J., Inamura,M., Asaka,M. and
 Hatakeyama,M.
 A novel human RSCAP-like gene that maps within the prostate cancer
 susceptibility locus at chromosome 1q25
 FEBS Lett. 441 (1), 127-131 (1998)
 99093006
 2 (bases 1 to 4368)
 Noto,S., Maeda,T., Hattori,S. and Hatakeyama,M.
 Direct Submission
 Submitted (11-PEB-1998) Department of Viral Oncology, The Cancer
 Institute, Japanese Foundation for Cancer Research, 1-37-1
 Kami-Ikebukuro, Toshima-ku, Tokyo, Japan
 Location/Qualifiers
 1..4368
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /map="1q24"
 126..3545
 /note="ras GTPase activating protein"
 /codon_start=1
 /product="nGAP"
 /protein_id="AA040481.1"
 /db_xref="GI:4105589"
 /translation="MQTPEVPAERSPRRRISGPTSTSEKPNMSMDTANTSPFKVPGFFS"

25916 TCTAGATCTCTCTGTCACACTGGGGCGGTGGCGCTGGCGCCCTGCATCAGT 25867

150 erileLeuGlyGlnclnuPheCysPheGluValThrThrSer...SerGly 165
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
25866 CGGTTCGGGACGTCGTCACTCACTTTTTTCCAGATACGGGGCGGCCCTCGTGCC 25817

166 ThrLysCysPheAlaCysArgSerAlaLaGluAraGlyAspLysTrpIleGlu 182
::: :::::::::::::: |||| | ||||| ||||| ||||| ||||| ||||| |||||

25816 GAGCGGTACTACTCATCGGATCGCGCACAGAGCCGACCCTTTGGATCTA 25767

182 uAsnLeuGlnAraGlaValLysProAsnLysAspAsnSerArgArgVala 199
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
25766 CTCGCTGCGCAAGTCGATCGCTCCGAATGCAGAGACACACGCGTCGCACGG 25717

199 sPaSnValLeuLysLeuTrpIleIleGluAlaArqGluLeuProPropLys 215
|||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
25716 ACAACTCGCTCAAGATGTGGTGTACGAGGCGAAANAATCTCCGCCCAAG 25667

216 LysArgTyrrTyrrCysGluLeuCysLeuAspMetLeuTyrrAlaArqTh 232
|||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

25666 AAGCGTTACTTTTTCGCACTGCAATTGGCAAGACGCTGTACGCCCGGAC 25617

232 rThrSerLysProArgSerAlaSerGlyAspThrValPheTrpGlyGluH 249
|::: ||| ::::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||::: ||:::
25616 TTCGDTGAAGATGCGACAGC..... GATCTGCTGTTTTTGGGGGGAGC 25576

249 IsPheGluPheAsnLeuProAlaValArqAlaLeuArgLeuHisLeu 265
::: ::::: |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
25575 ACTTCGATTTCCCGGCACATACCGAGATTAAATGTATCATCTGTTAAAGTT 25526

266 TyrArgAspSerAspLysArgLysLysAspLysAlaGlyTrpValGl 282
::: ||||| ::: ||||| ::: ||||| ::: ||||| ::: ||||| ::: ||||| ::: |||||
25525 TTCCGTGAGGTGGACAAGAAGAAGCGGGCAATAACCAANTTTGTGGG 25476

282 yLeuValThrValProValAlaThrLeuAlaGlyArgHisPheThrGluG 299
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
25475 ATCCGTGAAGATACCCGTGCACGATGTCACTCCAGATPGCCCTGCCAAC 25426

299 InTrpTyrrProValThrLeuProThrGlySerGlyGlyserGlyGlyMet 315
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
25425 AATGTAATCCCATA...CTGAGCGACAAGCGGCGCAGACTCTGGCAGG 25379

316 GlySerGlyGlyGlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlyGly 332
|||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
25378 ACATCGGC.....GGCGGCGCAGTGGGTCCAAGACNA 25344

332 sGlyGlyCysProAlaValArgLeuLysAlaArgTyrrGlnThrMetSerl 349
|||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
25343 GGAGCAATTCGCCACCGCTGAGGATCAAGTGTCTTCACAGACACCGACA 25294

349 LeuPrometGluLeuTyrrLysGluPheAlaGluTyrrValThrAsnHis 365
|||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
25293 TCCTGCCCATCAATGTGTACGGCAACTTTTTCGACGTACCTCAAGGAGAAC 25244

366 TyrArgMetLeuCysAlaValLeuGluProAlaLeuAsnValLysGlyLy 382
|||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
25243 TATAGCCGCTGTGCGACCCCTGGAGCGGTGATPCGGAGTCAGGCCCAA 25194

382 sGluGluValAlaSerAlaLeuValHisIleLeuGlnSerThrGlyLysa 399
|||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
25193 GGAGGACATTCGACAGGCACGTGGTCTCTGATGCACGACAGGGATTGG 25144

399 LaLysAspPheLeuSerAspMetAlaMetSerGluValAspArgPheMet 415
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
25143 CGGGCGCTTCTCCACCATGTGGTGGCCCTCGATCTCTCGAGATTGGC 25094

416 GluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThrLysAlaIl 432
::: ::::: |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
25093 GAT...CAGAGGCTTACGTTTCAGGGGCAACTCCTTGGCCACCAAGAGCAT 25047

432 eGluGluTyrrMetArgLeuIlLeGlyGlnLysTyrrLeuLysAspAlaIlLeG 449
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
25046 GGAGCAATTCCTCAAGCTGAGCGGCAACAGTATCTCGACAGACACATPAT 24997

[illegible]

| | | | | | |
|----------------------------|---|------------|-----|-----|-------------|
| 22785 | TTGAGGCTCAACGGCTCGC.....TCAGCGCGCGCAGTAGCTCTCTCC | 22745 | | | |
| 1256 | oProGln.....ProHisP | 1261 | | | |
| 22744 | ACCTCAGCATGCTGGCTCCACCTCGAATCTTTGGCAAGCGCGGCTCATCC | 22695 | | | |
| 1261 | roProHisProGlyCysArgSerGlnArgThrAlaSerSerGlyThrPro | 1277 | | | |
| 22694 | CGCTACGGCGGCTGAATGGCGCGCTGAAGCGCGGAGGATGCTACGGCA | 22645 | | | |
| 1278 | GlnThrThrSerProPro.....SerIleThrAspLeuLeuPr | 1290 | | | |
| 22644 | GTGGCTACGGCGCGCAGCGAGCTGGAAATGTGGCTATGGCTGTGCCACTTCC | 22595 | | | |
| 1290 | o.Cys.....AlaProTyrProGlyProProSerValThrAspLeuLeu | 1304 | | | |
| 22594 | AGTGCGGGGGGACCATCAACATCCCAACGACGACGACGAATCCGAT | 22545 | | | |
| 1305 | ProSerAlaProAspProGlyThrSerProThrThrArgThr..... | 1318 | | | |
| 22544 | GCAGCAGCAGCAGCAAGGACGGGATCAGGACCAACAGCAGTATCGG | 22495 | | | |
| 1319 |GlyArgHisGlnGlyThrAlaGlyLeuSerProLeuThrProp | 1333 | | | |
| 22494 | GCAGTGTGGCGGCGAGCGTGGGATCGGCACATCAGCGGCTCAGAGCGGC | 22445 | | | |
| 1333 | ro.....TrpGlyThrHis | 1337 | | | |
| 22444 | CTGAGCTTGGACTCGGCGGCAC | 22422 | | | |
| seq_name: gb_htg4:AC012162 | | | | | |
| seq_documentation_block: | | | | | |
| LOCUS | AC012162 | 191504 bp | DNA | HTG | 09-FEB-2000 |
| DEFINITION | Drosophila melanogaster chromosome X clone BACROIN10 (D115) | | | | |
| | RC1-98.01.N10 map 16D-16F strain Y; cn bw sp, *** SEQUENCING IN | | | | |
| | PROGRESS ***, 142 unordered pieces. | | | | |
| ACCESSION | AC012162 | | | | |
| VERSION | AC012162.9 | GI:6957590 | | | |
| KEYWORDS | HTG; HTGS_PHASE1. | | | | |
| SOURCE | fruit fly. | | | | |
| ORGANISM | Drosophila melanogaster | | | | |
| REFERENCE | Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhorf,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hinkle,A., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Lee,B., Lewis,S., Li,P., Ling,H., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Richards,S., Sethi,H., Svirskas,R.R., Wan,K.H., Webster,D., Woolley,P., Yang,S., Yee,M., Yu,C. and Rubin,G.M. | | | | |
| AUTHORS | Sequencing of Drosophila melanogaster | | | | |
| | Unpublished | | | | |
| TITLE | 2. (bases 1 to 191504) | | | | |
| JOURNAL | Celniker,S.E., Agbayani,A., Arcaina,T.T., Baxter,E., Blazej,R.G., Butenhorf,C., Champe,M., Chavez,C., Chew,M., Ciesiolka,L., Doyle,C.M., Farfan,D.E., Galle,R., George,R.A., Harris,N.L., Hoskins,R.A., Houston,K.A., Hummasti,S.R., Karra,K., Kearney,L., Kim,E., Lee,B., Lewis,S., Li,P., Lomotan,M.A., Mazda,P., Moshrefi,A.R., Moshrefi,M., Nixon,K., Pacleb,J.M., Park,S., Pfeiffer,B., Poon,L., Sequeira,A., Sethi,H., Svirskas,R.R., Wan,K.H., Weinburg,T., Zhang,R., Zieran,L.L. and Rubin,G.M. | | | | |
| REFERENCE | Direct Submission | | | | |
| AUTHORS | Submitted (21-OCT-1999) Drosophila Genome Center, Lawrence Berkeley Laboratory, MS 64-121, Berkeley, CA 94720, USA | | | | |
| COMMENT | On Feb 11, 2000 this sequence version replaced gi:5838825. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive web site (http://www.fruitfly.org/sequence/) or send email | | | | |

to bdqpfuitfly.berkeley.edu. All contigs in this submission meet the following cutoffs: length >= 200 bases.

* NOTE: This is a 'working draft' sequence. It currently consists of 142 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

| | | | |
|---|--------|---------|-----------------------------|
| * | 1 | 456: | contig of 456 bp in length |
| * | 457 | 536: | gap of unknown length |
| * | 537 | 1004: | contig of 468 bp in length |
| * | 1005 | 1084: | gap of unknown length |
| * | 1085 | 1958: | contig of 874 bp in length |
| * | 1959 | 2038: | gap of unknown length |
| * | 2039 | 2248: | contig of 210 bp in length |
| * | 2249 | 2288: | gap of unknown length |
| * | 2329 | 2858: | contig of 530 bp in length |
| * | 2859 | 2938: | gap of unknown length |
| * | 2939 | 3325: | contig of 387 bp in length |
| * | 3326 | 3405: | gap of unknown length |
| * | 3406 | 4165: | contig of 760 bp in length |
| * | 4166 | 4245: | gap of unknown length |
| * | 4246 | 4716: | contig of 471 bp in length |
| * | 4717 | 4796: | gap of unknown length |
| * | 4797 | 5373: | contig of 577 bp in length |
| * | 5374 | 5453: | gap of unknown length |
| * | 5454 | 6050: | contig of 597 bp in length |
| * | 6051 | 6130: | gap of unknown length |
| * | 6131 | 6720: | contig of 590 bp in length |
| * | 6721 | 6800: | gap of unknown length |
| * | 6801 | 7461: | contig of 661 bp in length |
| * | 7462 | 7541: | gap of unknown length |
| * | 7542 | 8104: | contig of 563 bp in length |
| * | 8105 | 8184: | gap of unknown length |
| * | 8185 | 8505: | contig of 321 bp in length |
| * | 8506 | 8585: | gap of unknown length |
| * | 8586 | 9227: | contig of 642 bp in length |
| * | 9228 | 9307: | gap of unknown length |
| * | 9308 | 10091: | contig of 784 bp in length |
| * | 10092 | 10171: | gap of unknown length |
| * | 10172 | 11099: | contig of 928 bp in length |
| * | 11100 | 11179: | gap of unknown length |
| * | 11180 | 11747: | contig of 568 bp in length |
| * | 11748 | 11827: | gap of unknown length |
| * | 11828 | 12767: | contig of 940 bp in length |
| * | 12768 | 12847: | gap of unknown length |
| * | 12848 | 13398: | contig of 551 bp in length |
| * | 13399 | 13478: | gap of unknown length |
| * | 13479 | 14007: | contig of 529 bp in length |
| * | 14008 | 14087: | gap of unknown length |
| * | 14088 | 14867: | contig of 780 bp in length |
| * | 14868 | 14947: | gap of unknown length |
| * | 14948 | 15562: | contig of 615 bp in length |
| * | 15563 | 15642: | gap of unknown length |
| * | 15643 | 16859: | contig of 1217 bp in length |
| * | 16860 | 16939: | gap of unknown length |
| * | 16940 | 17986: | contig of 1047 bp in length |
| * | 17987 | 18066: | gap of unknown length |
| * | 18067 | 19341: | contig of 1275 bp in length |
| * | 19342 | 19421: | gap of unknown length |
| * | 19422 | 20926: | contig of 1505 bp in length |
| * | 20927 | 21006: | gap of unknown length |
| * | 21007 | 22153: | contig of 1147 bp in length |
| * | 22154 | 22233: | gap of unknown length |
| * | 22234 | 23236: | contig of 1003 bp in length |
| * | 23237 | 23316: | gap of unknown length |
| * | 23317 | 24607: | contig of 1291 bp in length |
| * | 24608 | 24687: | gap of unknown length |
| * | 24688 | 25620: | contig of 933 bp in length |
| * | 25621 | 25700: | gap of unknown length |
| * | 25701 | 26300: | contig of 600 bp in length |
| * | 26301 | 26380: | gap of unknown length |
| * | 26381 | 27488: | contig of 1108 bp in length |
| * | 27489 | 27568: | gap of unknown length |
| * | 27569 | 28215: | contig of 647 bp in length |
| * | 28216 | 28295: | gap of unknown length |
| * | 28296 | 29440: | contig of 1145 bp in length |
| * | 29441 | 29520: | gap of unknown length |
| * | 29521 | 30587: | contig of 1067 bp in length |
| * | 30588 | 30667: | gap of unknown length |
| * | 30668 | 32042: | contig of 1375 bp in length |
| * | 32043 | 32122: | gap of unknown length |
| * | 32123 | 33740: | contig of 1618 bp in length |
| * | 33741 | 33820: | gap of unknown length |
| * | 33821 | 35375: | contig of 1555 bp in length |
| * | 35376 | 36258: | contig of 803 bp in length |
| * | 36259 | 36338: | gap of unknown length |
| * | 36339 | 38003: | contig of 1665 bp in length |
| * | 38004 | 38083: | gap of unknown length |
| * | 38084 | 39509: | contig of 1426 bp in length |
| * | 39510 | 39589: | gap of unknown length |
| * | 39590 | 40719: | contig of 1130 bp in length |
| * | 40720 | 40799: | gap of unknown length |
| * | 40800 | 42453: | contig of 1654 bp in length |
| * | 42454 | 42533: | gap of unknown length |
| * | 42534 | 43841: | contig of 1308 bp in length |
| * | 43842 | 43921: | gap of unknown length |
| * | 43922 | 45410: | contig of 1489 bp in length |
| * | 45411 | 45490: | gap of unknown length |
| * | 45491 | 46782: | contig of 1292 bp in length |
| * | 46783 | 46862: | gap of unknown length |
| * | 46863 | 48570: | contig of 1708 bp in length |
| * | 48571 | 48650: | gap of unknown length |
| * | 48651 | 51042: | contig of 2392 bp in length |
| * | 51043 | 51122: | gap of unknown length |
| * | 51123 | 52760: | contig of 1638 bp in length |
| * | 52761 | 52840: | gap of unknown length |
| * | 52841 | 54712: | contig of 1872 bp in length |
| * | 54713 | 54792: | gap of unknown length |
| * | 54793 | 56521: | contig of 1729 bp in length |
| * | 56522 | 56601: | gap of unknown length |
| * | 56602 | 58938: | contig of 2337 bp in length |
| * | 58939 | 59018: | gap of unknown length |
| * | 59019 | 60837: | contig of 1819 bp in length |
| * | 60838 | 60917: | gap of unknown length |
| * | 60918 | 62988: | contig of 2071 bp in length |
| * | 62989 | 63068: | gap of unknown length |
| * | 63069 | 65583: | contig of 2515 bp in length |
| * | 65584 | 65863: | gap of unknown length |
| * | 65864 | 68774: | contig of 3111 bp in length |
| * | 68775 | 68854: | gap of unknown length |
| * | 68855 | 71235: | contig of 2381 bp in length |
| * | 71236 | 71315: | gap of unknown length |
| * | 71316 | 74193: | contig of 2798 bp in length |
| * | 74194 | 74194: | gap of unknown length |
| * | 74194 | 77405: | contig of 3212 bp in length |
| * | 77406 | 77485: | gap of unknown length |
| * | 77486 | 80936: | contig of 3451 bp in length |
| * | 80937 | 81016: | gap of unknown length |
| * | 81017 | 83357: | contig of 2341 bp in length |
| * | 83358 | 84337: | gap of unknown length |
| * | 84338 | 87456: | contig of 4019 bp in length |
| * | 87457 | 87536: | gap of unknown length |
| * | 87537 | 91532: | contig of 3996 bp in length |
| * | 91533 | 91612: | gap of unknown length |
| * | 91613 | 97143: | contig of 5531 bp in length |
| * | 97144 | 97223: | gap of unknown length |
| * | 97224 | 105593: | contig of 8370 bp in length |
| * | 105594 | 105673: | gap of unknown length |
| * | 105674 | 112071: | contig of 6398 bp in length |
| * | 112072 | 112151: | gap of unknown length |
| * | 112152 | 121011: | contig of 8860 bp in length |
| * | 121012 | 121091: | gap of unknown length |


```

114219 CCCACACAGAGAGGGTTCGTGGCGGGTTCCCTGCACGACAGAGGC 114268
479 aAsnLeuArgMetCysCysGluLeuAlaLeuCysLysValValAsnSerH 496
114269 CGCCTTGGCTGGCGGGTCCGAGGGGGTGGCAGTGCATCTCCGAATGCG 114318
496 iCysValPheProArgGluLeuLysGluValPheAlaSerTrpArgLeu 512
114319 ACAAGCATTTCCGCCCCAGTTAGAAATTCCTTCGGAGCTCCGGAG 114368
513 ArgCysAlaGluArgGlyArgGluAspIleAlaAspArgIleSerAl 529
114369 CGCTTGCAGCAGCTGGCGCGTTCAGGATATGGCCGCAACCTGATCTCGG 114418
529 aSerLeuPheLeuArgPheLeuCysProAlaIleMetSerProSerLeuP 546
114419 GAGCATTTTCCTGGCTTCCTGTGGCGGGCCATCTGTGCGCGTCTCT 114468
546 heGlyLeuMetGlnGluTrpProAspGluGlnThrSerArgThrLeuThr 562
114469 TCATATCACCAGGGAACCTGGCGTCCGACGGGCCACCGCAATCTCACA 114518
563 LeuIleAlaLysValIleGlnAsnLeuAlaAsnPheSerLysPheThrSe 579
114519 CTGGTGGCCAGACCCCTGCAACATTTGCCAATTTCCACCGCTTCCAGGG 114568
579 rLysGluAspPheLeuGlyPheMetAsnGluPheLeuGluLeuLutrpG 596
114569 CAAAGAGAACTTATGGAGTTCTCAACGATTTCTCCTCGAGCGGCGG 114618
596 lysSerMetGlnGlnPheLeuTyroGluIleSerAsnLeu..... 608
114619 CTCGCATCAACACAGTTCTGGAGATTATATCCACGGCGGCGGAGCACCA 114668
609 .....AspThrLeuThrAsnSerSerPheGluGlyTyroIleAspLe 623
114669 GCTCCAGACTCGATCTCGAT.....TGGCGCGGTATCATCGACCA 114709
623 uGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGluValLeuProG 640
114710 GGGCAACAGTTGTCTACATACACAGTTGCTCAGGAAAGCTGGCCA 114759
640 lnLeuSerLysGluAlaLeuLeuLysLeuGlyProLeuProArgLeuLeu 656
114760 AGCTGCCGAGGCCAGGAGCAGCAGCTGGATCGTTGCAGCATATCTC 114809
657 SerAspIleSer.....ThrAlaLeu..... 663
114810 GATGAATCAGCGAGCCAAAGACATGGCATGGGCACAGCATCGCGGG 114859
663 ..... 663
114860 TGGATATTTCGGGCCAACCCTGCTACGCACTGATAGCAGCGGAATC 114909
664 .....ArgAsnPro.....AsnIle 668
114910 AGGAGAATCGCAATCGGGATCTCGGCTCGCAGCTGCGCTCCAATCG 114959
669 GlnArgGlnProSerArgGlnSerGluArgAlaArgSerGlnProMetVa 685
114960 GAGCAGTTACTGCCACAACAAGAGCGAGTTGGCCCGCAGCGCATCGGAT 115009
685 lLeuArgGlyPro....SerAlaGlu.....MetGlnGly.TyrMet 697
115010 TGTAGTAACCAATTTATCTGGCGGCGCGGCATCATCGAGGAGTACTTA 115059
698 MetArgAspLeuAsnSerSerIleAspLeuGlnSerPheMetAlaArgG 714
115060 CGCCGAATTTCTCTGGAGAGAGATATCTTTAGATACAAATGATCCCGCT 115109
714 yLeuAsnSerSerMetAspMetAlaArgLeuProSerProThrLysGluL 731

```

```

115110 AATG..... 115113
731 ysProProProProGlyGlyLysAspLeuPheTyrValSer 747
115113 ..... 115113
748 ArgProProLeuAlaArgSerSerProAlaTyrCysThrSerSerSerAs 764
115114 .....GCTTACTGCAGCAGCAGCAGCAG 115136
764 pIleThrClnProGluGlnLysMetLeuSerValAsnLysSerValSerM 781
115137 CAGCAGAGC.....AGCAGCAACAACAGCAACAGCA 115168
781 etLeuAspLeuGlnGlnLysPglyArgLeuAsnSerSerSer 797
115169 ACATC.....AGCAGCTGCAACAGCATGGCC 115194
798 ValSerAsn.....LeuAlaAlaValGly..AspLeuLeuHisSerSer 811
115195 ATCAGCAACACCCGACACAGCATCCACTCCAGATGCTCTCCAATTCA 115244
812 GlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaProAlaGlyAr 828
115245 CAACCTCCATT...GCGGCAACCAATATATGAGTTCCGCCAGGAGGCT 115291
828 gLeuSerGlnGlySerGlySerSerIleThrAlaAlaGlyMetArgLeu 845
115292 GCAGCATGCCCAATCGCAGCTCGATGGGCTCTCATCGCTT..... 115334
845 erGlnMetGlyValThrAspGlyValProAlaGlnGlnLeuArgIle 861
115334 ..... 115334
862 ProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaAspGlyProG 878
115334 ..... 115334
878 yProProAlaGlyHisGlyGlySerSerGly.....HisGlyProProS 893
115335 .....AATGGGAGCAGCAGCANTTTGCTGCAGCGC..... 115364
893 erSerHisHisHisHisHisHis.....HisHisHisArg 905
115365 .....CACCAGCAGCATGCCATCACCCGCAACTGTCATCCATCAC 115409
906 GlyGlyGluProProGlyAspThrPheAlaProPheHisGlyTyrSerLy 922
115410 TGC.....CGCGCG..... 115418
922 sSerGluAspLeuSerThrGlyValProLysProProAlaAlaSerIleL 939
115419 .....GCGCCACAGCAGCAGTGCCTCCAGCAGCTA 115446
939 euHis.....SerHisSerTyrSerAspGluPheGlyPro 950
115447 TGGAGCGCATGGATCGCATGAATATCCGTATATGTGCGCAC..... 115487
951 serGlyThrAspPheThrArgGlnLeuSerLeuGlnAspAsnLeuG 967
115488 AATGGCATGACTAGCAG..... 115505
967 nHisMetLeuSerProGlnIleThrIleGlyProGlnArgProAlaP 984
115506 .....ACCAGCAGCGCTTCGAGCAGCTCCAGGACACTGCCCGGGA 115548
984 roSerGlyPro.....GlyGlyGlySerGlyGlySerGlyGly 998
115549 ATGGAATCCCAATGCCCAATGCCAAGCTGGGAGCAGCAGCAACAATAACCAG 115598
999 GlyGlyGlyGlnProProProLeuGlnArgGlyLysSerGlnGlnLeu. 1014
115599 AGCGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 115648

```

```

1015 .....ThrValSerAlaAlaGlnLysProArgPro..... 1024
115649 GTTCGATACGAGCAGTCTGTCTGCAAGTCGCCACACCCATCATGA 115698
1025 .....Ser.SerGlyAsnLeuLeuGlnSerProGluProSerT 1037
115699 ATTCAGTTTGGGACCGCGGGCGGACGAGTCAACAACTGAAT 115748
1037 yGlyProAlaArgProArg.....Gln 1044
115749 TTGGGAATACCCGATCACTACAGTGGCTATGCGGGGTAATAATTGAA 115798
1045 GlnSerLeuSerLysGluLysSerIleGlySerGlySer..... 1059
115799 TCCTCAACTGGAATATGCCAGAACTTGGAGATCTGGACGATCTGTCA 115848
1060 .....GlyGlyGlyGlyGlyGlyGlyGlyLeuLysProSerIleThrL 1073
115849 AGTACGCCGAGGAGCATGACGTGGCGGACCAACGACCAACATCAATCAC 115898
1073 yGlnHisSer....GlnThrProSerThr.LeuAsnPro....ThrMetPr 1087
115899 AACCAGGTCAGCAGACACCGAGGACATCTGAAGCGCGCGCTTCC 115948
1087 oAlaSerGluArgThrValAlaTrpValSerAsnMetProHisLeuSerA 1104
115949 CGCAAGGAGCAGCTGTGGCGGAAAGCAGT.....CACTGCAGTT 115989
1104 laAspIleLysSerAlaHisIleGluArgGluGluLysLeuLysGlu 1120
115990 CTGGTACGAGCATCTCCACAAATCCG..... 116018
1121 TyrSerLysSerMetAspGluSerLeuLeuAspArgValLysGluTrpG 1137
116019 ...TCGCCTCGAGTCTCCGTCGCGTCCGTCGAG..... 116048
1137 uGluGluLysSerLeuLysGluArgLeuHisMetSerAsnArgLysL 1154
116049 .....AGCCAGCTGAAGCGCGCGATGGGAGTCAACAAATGCGCGCG 116088
1154 euGluGluTrpGluArgLeuLeuSerGlnGluGluGlnThrSerLys 1170
116089 TG...GCCTTCAAGAACTCCCTCTATCATGCTTCAGCCCAACAGTCTCG 116135
1171 IleLeuMetGlnTrpGlnAlaArgLeuGluGlnSerGluLysArgLeuR 1187
116136 TCAGATCATCGGCACAGAGTAATACACACGACGACGACGACGACGAC 116185
1187 gGlnGlnGlnValGluLysAspSerGlnIleLysSerIleIleGlyArgL 1204
116186 ACAACAACAGTTT.....GGCAGCGCTTGAACCAATTTGGCGCGGAC 116229
1204 euMetLeuValGluGluLeuArgArgAspHisProAlaMet..AlaG 1220
116230 TGGTGGCGCGAGCGCGCTTCTCTCAACAGTGGCGGAGCTTGGAGCGG 116279
1220 uProLeu.....ProGluProLys..... 1226
116280 GCCACTTTGACGCCCGAGCTCTCGGACGAACAGCTCTCGGCGGATAATTA 116329
1227 .....LysArgLeuLeuAspAlaGlnArgGlySerPheProPro 1239
116330 CTTCAGTTATGACGGCTGCGAGCTGTGGAGCGGGTATTGGCACCAAT 116379
1240 TrpValGlnGlnThrArgValArgTrpProHisLeuGlyThrAlaTrpPr 1256
116380 TGGAGCTCAACGCTCG.....TCAGCGCGCGGACGAGTCTCTCTCC 116420
1256 oProGln.....ProHisP 1261
116421 ACCTCAGCATCTGGCTCCACCTCGAATCTTGGCAAGAGCGCGGCTCATC 116470

```

1261 roProHisProGlyCysArgSerGlnArgThrAlaSerSerGlyThrPro 1277
116471 CGCTACGCGCGCTGAATGGCGCGCTGAAGCGGAGGATGCTACGGCA 116520
1278 GlnThrThrSerPro.....SerIleThrAspLeuLeuPr 1290
116521 GTGGCTACGCGCGCGGACGAGTGAATAATGGGCTATGGCTGTCCACTTCC 116570
1290 o.Cys.....AlaProTrpProGlyProProSerValThrAspLeuLeu 1304
116571 AGTGGCGCGGACGACATCAACATCCACGAGCAGCAGGATCCGAT 116620
1305 ProSerAlaProAspProGlyThrSerProThrThrArgThr..... 1318
116621 GCAGCAGCAGCAGCAAGGAGGATCAGCAACACAGCAGTATGCCG 116670
1319GlyArgHisGlnGlyThrAlaGlyLeuSerProLeuThrProp 1333
116671 GCAGTGGCGCGGACGCTGGGATCGGCACATCAGCGGCTCAGAGCGC 116720
1333 ro.....TrpGlyThrHis 1337
116721 CTGAGCTTGGAGCTCGCGCGCAC 116743
seq_name: gb_inl:AE003506
seq_documentation_block:
LOCUS AE003506 300994 bp DNA INV 04-OCT-2000
DEFINITION Drosophila melanogaster genomic scaffold 14200001386053 section 23
of 30, complete sequence.
ACCESSION AE003506 AE002593
VERSION AE003506.1 GI:7293355
KEYWORDS HTS.
SOURCE fruit fly.
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS
Adams,M.D., Celisner,S.E., Holt,R.A., Evans,C.A., Gocayne,J.D.,
Amanatides,P.G., Scherer,S.E., Li,P.W., Hoskins,R.A., Galie,R.F.,
George,R.A., Lewis,S.E., Richards,S., Ashburner,M., Henderson,S.N.,
Sutton,G.G., Wortman,J.E., Yandell,M.D., Zhang,O., Chen,L.X.,
Brandon,R.C., Rogers,Y.H., Blazej,R.G., Champe,M., Pfeiffer,B.D.,
Wan,K.H., Doyle,C., Baxter,E.G., Helt,G., Nelson,C.R., Gabor
Miklos,G.L., Abril,J.F., Agbayani,A., An,H.J.,
Andrews-Frannkoch,C., Baldwin,D., Ballew,R.M., Basu,A.,
Baxendale,J., Bayraktaroglu,L., Beasley,E.M., Beeson,K.Y.,
Benos,P.V., Berman,B.P., Bhandari,D., Bolshakov,S., Borkova,D.,
Botchan,M.R., Bouck,J., Brokstein,P., Brottier,P., Burtis,K.C.,
Busam,D.A., Butler,H., Cadieu,E., Center,A., Chandra,I.,
Cherry,J.M., Cawley,S., Dahlke,C., Davenport,L.B., Davies,P., de
Pablos,B., Delcher,A., Deng,Z., Mays,A.D., Dew,I., Dietz,S.M.,
Dodson,K., Doup,L.E., Downes,M., Dugan-Rocha,S., Dunkov,B.C.,
Dunn,P., Durbin,K.J., Evangelista,C.C., Ferraz,C., Ferreira,S.,
Fleischmann,W., Folsler,C., Glodek,A., Gong,F., Garg,N.S.,
Gelbart,W.M., Glasser,K., Glodzik,A., Gong,F., Garg,N.S.,
Guan,P., Harris,M., Harris,N.L., Harvey,D., Heiman,T.J.,
Hernandez,J.R., Houck,J., Hostin,D., Houston,K.A., Howland,T.J.,
Wei,M.H., Ibegwam,C., Jalili,M., Kalush,F., Karpen,G.H., Ke,Z.,
Kenshin,J.A., Ketchum,K.A., Kimmel,B.E., Kodira,C.D., Kraft,C.,
Krivitz,S., Kulp,D., Lai,Z., Lasko,P., Lei,Y., Levitsky,A.A.,
Li,J., Li,Z., Liang,Y., Lin,X., Liu,X., Mattei,P., McIntosh,T.C.,
McLeod,M.P., McPherson,D., Merkulov,G., Milshina,N.V., Mobarry,C.,
Morris,J., Moshrefi,A., Mount,S.M., Moy,M., Murphy,B., Murphy,L.,
Muzny,D.M., Nelson,D.B., Nelson,D.R., Nelson,K.A., Nixon,K.,
Nusskern,D.R., Pacle,J.M., Palazzolo,M., Pittman,G.S., Pan,S.,
Pollard,J., Puri,V., Reese,M.G., Reinert,K., Remington,K.,
Saunders,R.D., Scheefer,P., Shen,H., Shue,B.C., Siden-Kiamos,I.,
Simpson,M., Skupski,M.P., Smith,T., Spier,E., Spradling,A.C.,
Stapleton,M., Strong,R., Sun,E., Svirskaas,R., Tector,R.,
Venter,E., Wang,A.H., Wang,Z.Y., Wang,Z.Y., Wassarman,D.A.,
Weinstock,G.M., Weissenbach,J., Williams,S.M., Woodage,T.,

Worley,K.C., Wu,D., Yang,S., Yao,Q.A., Ye,J., Yeh,R.F.,
Zaveri,J.S., Zhan,M., Zhang,G., Zhao,Q., Zheng,L., Zheng,X.H.,
Zhong,F.N., Zhou,W., Zhou,X., Zhu,S., Zhu,X., Smith,H.O.,
Gibbs,R.A., Myers,E.W., Rubin,G.M. and Venter,J.C.
The genome sequence of *Drosophila melanogaster*
Science 287 (5461), 2185-2195 (2000)
20196006
2 (bases 1 to 300994)
Adams,M.D., Celniker,S.E., Gibbs,R.A., Rubin,G.M. and Venter,C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA

FEATURES
Location/Qualifiers
1..300994
/organism="Drosophila melanogaster"
/chromosome="X"
/db_xref="taxon:7227"
complement(join(<16198..16317,16386..16574,16635..16845,
16915..17053,17129..17411,17526..17607,17676..17975,
18107..>18483))
/gene="CG8527"
/product="CT24901"
/db_xref="FLXBASE:FBan0008527"
/db_xref="FLXBASE:FBgn0030844"
complement(<16198..>18483)
/gene="CG8527"
/map="16B4-16B6"
/db_xref="FLXBASE:FBan0008527"
/db_xref="FLXBASE:FBgn0030844"
complement(join(16198..16317,16386..16574,16635..16845,
16915..17053,17129..17411,17526..17607,17676..17975,
18107..18483))
/gene="CG8527"
/note="CG8527 gene product"
/codon_start=1
/db_xref="FLXBASE:FBan0008527"
/db_xref="FLXBASE:FBgn0030844"
/protein_id="AAF48734.1"
/db_xref="GI:7293356"
/translation="MPOEKRPPESTORFLETLVIFRSLIYOTKEFFONSLHGVY
VIASGRPIGEKFWCFSTSIGAVTALVIMSLWEKFTQNTPIGLDTRFNQNVSP
TIVCPAEAFDHTYEKVTINTLYANIDEAQAQYTFKRLILSLNFENVRAKQVLSQ
IPQNLDAHTIREWAFGLHEDYDKNVEVSKYDEIDPCDHPEPIYTHGFCYAFNSR
FKSPTEDTICAGFDHETDKWALFFIPNSTRIFFISNEEYFGDFNAQIDWSEP
QLVEVRSKNTYTDARQLSICQKRCIFSEVILNYPDAYTFSCMKQCRNKA
KLCNRPFPYKPIPNVPMCSIKDFDCIDDEFKNTINIKDCLQCELSCKTWFNDKLI
KMSDRPSGLVLEFLTWPIIRYKREVLFGWDLVIFGGIASLFGFSLSGVEIIV
YFTLRACMVKNQELYEIEEKIROBPPKIDLLSLKSHNPRIGDPTSAVLKVP
AEESSGNPLGGYQKHEKDYSLNTKNYKTPKVPDYGYTSKTKNSWTTIDHSQYWP
"
complement(join(21273..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25449))
/gene="CG8465"
/product="CT24743"
/db_xref="FLXBASE:FBan0008465"
/db_xref="FLXBASE:FBgn0030845"
complement(join(21273..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28246))
/gene="CG8465"
/product="CT24757"
/db_xref="FLXBASE:FBan0008465"
/db_xref="FLXBASE:FBgn0030845"
complement(join(21273..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28246))
/gene="CG8465"
/product="CT24733"
/db_xref="FLXBASE:FBan0008465"
/db_xref="FLXBASE:FBgn0030845"
complement(<21273..>28555)
/gene="CG8465"

REFERENCE
TITLE
JOURNAL
MEDLINE
AUTHORS
DIRECT SUBMISSION
JOURNAL
JOURNAL

CDS
/map="16B4-16B7"
/db_xref="FLXBASE:FBan0008465"
/db_xref="FLXBASE:FBgn0030845"
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28092))
/gene="CG8465"
/note="CG8465 gene product [alt 1]"
/codon_start=1
/db_xref="FLXBASE:FBan0008465"
/db_xref="FLXBASE:FBgn0030845"
/protein_id="AAF48735.1"
/db_xref="GI:7293357"
/translation="MSTYFGVYIPTSKACGFCGVSQCIGSIAAVNIKPSNPASGSAS
VASGSPGSAASVOTGNADGSAATKIEDPDPDPLMFIETKSKALDILRHKEA
RLREFPNLEQAESYVOFGFESIALKRFCKAPESKPIITISGSKSPSTSDNSCS
SSPTGNSGFTIPLGNSNSNLLSDSPTSSSSSNVIANGROQOQOQOQOQOQ
PDVSGEGPPFRAPTQBLVFRKQIBRGHIDRVKRIWENPRFLISSGDTLSKEGC
RYNAHICAQVKNARIQAQLLKTISDREFTQLYVKGKSGKMAALISLDIYLNMP
DKGRGTEPLHFAAKNGHVAEVLVSPCKSLRNHGEKPEKIEICLNANATHVITIK
KLELLYDHPHVPVLRQSNTLPKVGQPSKDPNLOHKAADDYEGISVDAISALA
GPMSEKAMNFIYRRKTPPRVSNVMSPLAGSPSPVKYTPSKSIFDRSAGNSPVH
SCRVLFSPLAEATSSPKTKNYPNGTNECHNNNVKPIYEPATPIRKMPPDLF
MAYNNNSFSPSLADDSQILDMSLSRNASLNSDFRHHIKNTDIEKGLDEVGRQL
ARQBLEWREYWDFLSDIDIGTEGLARLEAYFLEKTEQADKSETVWNAHLHOYF
DSMAGEQQOLRKDKNEAGATSPSAGVMTPTCYEKSQVFAKRTITLKNKIGMY
SINDTLICELRLKLSLVSPKDDARFISVDPSKVSIRAHVLSVTSQVSVYAMRL
QILMLRSLROLLADERGRHGLGCVCALLMLBQAPTSVHLDPDTLKTBEELCAAW
ETEOCCARVNASASKAWRQOSDDYDQEQVLPFDCTNVTLPYSGSSSEDENFRT
PQSLSGISMDLEPRYELFIFGNEPTKROLDVLNALSNDIDKETLPHVYAWKTAME
SYSCAENLNKYVKQPEPWYSSTSSNSQPLLHKLRLIATPKLNAVYSGRRSGGPI
APVTPRLARTSAASIOVASETNGESVGTAVTPASPILSFAALTAAVTSOTPLNKVR
GLFSQYDQRSYNEGDTPLGNRN"
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25370,27861..28092))
/gene="CG8465"
/note="CG8465 gene product [alt 2]"
/codon_start=1
/db_xref="FLXBASE:FBan0008465"
/db_xref="FLXBASE:FBgn0030845"
/protein_id="AAF48736.1"
/db_xref="GI:7293358"
/translation="MSTYFGVYIPTSKACGFCGVSQCIGSIAAVNIKPSNPASGSAS
VASGSPGSAASVOTGNADGSAATKIEDPDPDPLMFIETKSKALDILRHKEA
RLREFPNLEQAESYVOFGFESIALKRFCKAPESKPIITISGSKSPSTSDNSCS
SSPTGNSGFTIPLGNSNSNLLSDSPTSSSSSNVIANGROQOQOQOQOQOQ
PDVSGEGPPFRAPTQBLVFRKQIBRGHIDRVKRIWENPRFLISSGDTLSKEGC
RYNAHICAQVKNARIQAQLLKTISDREFTQLYVKGKSGKMAALISLDIYLNMP
DKGRGTEPLHFAAKNGHVAEVLVSPCKSLRNHGEKPEKIEICLNANATHVITIK
KLELLYDHPHVPVLRQSNTLPKVGQPSKDPNLOHKAADDYEGISVDAISALA
GPMSEKAMNFIYRRKTPPRVSNVMSPLAGSPSPVKYTPSKSIFDRSAGNSPVH
SCRVLFSPLAEATSSPKTKNYPNGTNECHNNNVKPIYEPATPIRKMPPDLF
MAYNNNSFSPSLADDSQILDMSLSRNASLNSDFRHHIKNTDIEKGLDEVGRQL
ARQBLEWREYWDFLSDIDIGTEGLARLEAYFLEKTEQADKSETVWNAHLHOYF
DSMAGEQQOLRKDKNEAGATSPSAGVMTPTCYEKSQVFAKRTITLKNKIGMY
SINDTLICELRLKLSLVSPKDDARFISVDPSKVSIRAHVLSVTSQVSVYAMRL
QILMLRSLROLLADERGRHGLGCVCALLMLBQAPTSVHLDPDTLKTBEELCAAW
ETEOCCARVNASASKAWRQOSDDYDQEQVLPFDCTNVTLPYSGSSSEDENFRT
PQSLSGISMDLEPRYELFIFGNEPTKROLDVLNALSNDIDKETLPHVYAWKTAME
SYSCAENLNKYVKQPEPWYSSTSSNSQPLLHKLRLIATPKLNAVYSGRRSGGPI
APVTPRLARTSAASIOVASETNGESVGTAVTPASPILSFAALTAAVTSOTPLNKVR
GLFSQYDQRSYNEGDTPLGNRN"
complement(join(21399..21792,21868..21999,22058..22191,
22257..22823,22889..23887,24168..24320,24398..25116,
25191..25410))
/gene="CG8465"
/note="CG8465 gene product [alt 3]"
/codon_start=1
/db_xref="FLXBASE:FBan0008465"
/db_xref="FLXBASE:FBgn0030845"

CDS

```
/protein_id="AAF48737.1"
/db_xref="GI:7293359"
/translation="MPTHQHRHDAADSLPLWLFTEKSKALDILRHFKARLREPP
NLQAEYVGFGEFESLALKRFCKAKPEKPIITISGGYKSSPTSTDNSSCSSTPCN
GSGFIPLIGSSSSMNLISDPTSPSSSNVLIANGQQOQOQOPOOPDVSGE
GPPRAPTKQELVEFKQIEGHIDRVKRIIWNPRFLISSGDTPTSLKEGCRYNAMH
ICQVKNARIQLULLATIISDREFTQYVGKSGRMCAALNISLDYILNPNPDGRGE
TPLHFAAKNGHVAMVEVLVYPECKSLRNHEGKEIKELIURNANATHVTIKKLELL
```

```
alignment_scores:
  Quality: 126.50      Length: 1512
  Ratio: 1.591        Gaps: 62
  Percent Similarity: 52.513  Percent Identity: 28.836

alignment_block:
US-09-294-298-4 x AE003506 ..

Align seg 1/1 to: AE003506 from: 1 to: 300994

11 ProValGlnValProCysProLeuLeuProThrAlaSer.....Le 24
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195596 CCCTACTAATGCCATATCTACGCTTCTTTCGACAGATCGCTCTAT 195645
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
24 uSerAlaAlaAlaLeu.....ProAlaAlaPheArgGlyA 37
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195646 GAGAGCGGTGGCGGTGGATGACGCGCCACCGCCGCTT..... 195688
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
37 snThrThrTrpValAlaArgGlyGlyArgValSerProGlyGlyLysGln 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195689 .....TGGGAGCAACACAGCAGCAGGACCAATGCCACCTCGC 195730
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
54 TyrSerMetGluAlaAlaProAlaAlaProPhe..... 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195731 GTTTCACCAACTCTTTTCCAGAGAAGTTCCTCCAGTATATATATATA 195780
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
65 .....ArgProSerGln.....GlyPheLeuSerA 73
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195781 TATATATTCAGGTTCACATATTAATAATATGCCATTTGTTTCTTGGCC 195830
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
73 rArgLeuLysSerSerLysArgThrLysSerGlnProLysLeuAsp 89
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195831 ACAGA...TCCAATCCTTTGAAGCGGACCAAGTCGGTGACCAAGCTGGAG 195877
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
90 ArgThrSerSerPheArgGlnLeuProArgPheArgSerAlaAspHI 106
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195878 CGGACCAAG..... 195886
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 sAspArgAlaArgLeuMetGlnSerPheLysGlnSerHisGln 123
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195887 .....CGCGGATCGCGGCGACTGAGGGCTCCCGCTCGCACGAGA 195926
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
123 erLeuLeuSerProSerSerAlaAlaGluAla..... 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195927 GTCTGTGTCCAGTACGCGCTATCTCCACCATAGGTGAGTCTCGGCTC 195976
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
133 ..... 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
195977 ATATGGAGTTTGGCCAAACCTCAAAACCAITTTCTTTTCTTCTCTCT 196026
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
134 LeuGluLeuAsnLeuAspGluAspSerLleLysProValHisSer 150
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196027 TCTAGATCTCTGCTACCTGGGCGGTGGCGGTGGCGCGCTGATCAGT 196076
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
150 erLleLeuGlyGlnGluPheCysPheGluValThrThrSer...SerGly 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196077 CGGTTCTGGAGCTGCTACTCTCTTCCAGGTACGGGCGCGCCCTCGTGGC 196126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
166 ThrLysCysPheAlaCysArgSerAlaAlaGluArgAspLysTrpIleG 182
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196127 GAGCGTACTACTCATCGATCGCGCCAGGAGCGGACCTTTGGATCTA 196176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
*182 uAsnLeuGlnArgAlaValLysProAsnLysAspAsnSerArgValA 199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
196177 CTGCTCGCGAAGTCGATCGCTCCGAATCGAGGACACAGCGCTCGCAGG 196226
199 spAsnValLeuLysLeuTrpIleGluAlaArgGluLeuProLys 215
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196227 ACAACTCGCTGAAGATCGGTGTAGACGCGGAAAAAATCTGCCGCCAAG 196276
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
216 LysArgTyrTrpCysGluLeuCysLeuAspMetLysAlaArgTh 232
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196277 AAGCGTTACTTTTGCACACTGCAATTGGCAAGACGCTACGCGCGGAC 196326
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
232 rThrSerLysProArgSerAlaSerGlyAspThrValPheTrpGlyGluH 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196327 TTCGTTGAAGCTGCAGAG.....GATCTGCTGTTTGGGGGGAGC 196367
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
249 isPheGluPheAsnAsnLeuProAlaValArgAlaLeuArgLeuHisLeu 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196368 ACTTCGATTTCCTCCGAGATTAATGTGATCACTGTTACGTT 196417
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
266 TyrArgAspSerAspLysArgLysLysAspLysAlaGlyTrpValG 282
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196418 TTCCGTGAGGTGACAAAGAAAGAGCGGACCAATACCAATTTGTGGG 196467
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
282 yLeuValThrValProValAlaThrLeuAlaGlyArgHisPheThrGluG 299
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196468 ATCGTGAAGATACCTGTCAGCATGTCACTCCAGATTGCCCTCGGAC 196517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 InTrpTyrProValThrLeuProThrGlySerGlySerGlyGlyMet 315
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196518 AATGTTATCCATA...CTGAGCGACAAGCGGCGGACAGTGTGGCAGG 196564
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
316 GlySerGlyGlyGlyGlySerGlyGlyGlySerGlyGlyLysGlyL 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196565 ACATCGGC.....GGCGGCGGAGTGGTCCCAAGGACAA 196599
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
332 sGlyGlyCysProAlaValArgLeuLysAlaArgTrpGlnThrMetSer 349
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196600 GGACAAATGCCACGCTGAGTCAAGTGTCTTTCCAGAGCACCAGACA 196649
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
349 leLeuProMetGluLeuTyrLysGluPheAlaGlyTrpValThrAsnHis 365
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196650 TCCTGCCATCAATGTGTACGGCACTTTTGTGACGTACCTCAAGGAGAA 196699
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
366 TyrArgMetLeuCysAlaValLeuGluProAlaLeuAsnValLysGly 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196700 TATAGCGGTGTGCGAGACCTGGAGCGGTGATCGGASTCAAGGCCNA 196749
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 sGluGluValAlaSerAlaLeuValHisLysLeuGlnSerThrGlyLys 399
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196750 GGAGGACATTTGGACGACCTGTTGCTGTGATGACGACGAGGATTGG 196799
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
399 laLysAspPheLeuSerAspMetAlaMetSerGluValAspArgPheMet 415
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196800 CGGCGCTTCTCCACGATGTGGTGGCCCTCGATCTGCTGCGAGTTGGC 196849
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
416 GluArgGluHisLeuLlePheArgGluAsnThrLeuAlaThrLysAla 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196850 GAT...CAGAGGCTTACGTTTCAGGGGCACTCTTGGCCACCAAGAGCAT 196896
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
432 eGluGlyTrpMetArgLeuLleGlyGlnLysTyrLeuLysAspAlaIleG 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196897 GGAGCATTCCTCAAGCTGACGGCGCAACAGTATCTGACGAGACACACT 196946
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
449 lyGluPheLleArgAlaLeuTyrGluSerGluGluAsnCysGluValAsp 465
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196947 CCGCACCATAAAGAGCTAATTCAGTCGGAGAGGAGTTCGCGAGTGGAT 196996
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
466 ProIleLys.....CysThrAlaSerSerLeuAlaGluHisGlnAl 479
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
196997 CCCACCAAGACGAGCGGTTCGTCGGCGGTTCGTCGACGACACAGCAGC 197046
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
479 aAsnLeuArgMetCysCysGluLeuAlaLeuCysLysValValAsnSerH 496
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
197047 CGCCTTGGTGGCGGTTCGCGAGGGCGTGGCAGTCATCTTCGATCGC 197096
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


* 590 1201: contig of 612 bp in length
* 1202 1281: gap of unknown length
* 1282 2394: contig of 1113 bp in length
* 2395 2474: gap of unknown length
* 2475 3512: contig of 1038 bp in length
* 3513 3593: gap of unknown length
* 3593 4500: contig of 908 bp in length
* 4501 4580: gap of unknown length
* 4581 5465: contig of 885 bp in length
* 5466 5545: gap of unknown length
* 5546 6253: contig of 708 bp in length
* 6254 6333: gap of unknown length
* 6334 7105: contig of 772 bp in length
* 7106 7185: gap of unknown length
* 7186 8077: contig of 892 bp in length
* 8078 8157: gap of unknown length
* 8158 9047: contig of 890 bp in length
* 9048 9127: gap of unknown length
* 9128 9530: contig of 403 bp in length
* 9531 9610: gap of unknown length
* 9611 10173: contig of 563 bp in length
* 10174 10253: gap of unknown length
* 10254 10841: contig of 588 bp in length
* 10842 10921: gap of unknown length
* 10922 11807: contig of 886 bp in length
* 11808 11887: gap of unknown length
* 11888 13165: contig of 1278 bp in length
* 13166 13245: gap of unknown length
* 13246 14526: contig of 1281 bp in length
* 14527 14606: gap of unknown length
* 14607 15801: contig of 1195 bp in length
* 15802 15881: gap of unknown length
* 15882 16640: contig of 759 bp in length
* 16641 16720: gap of unknown length
* 16721 17679: contig of 959 bp in length
* 17680 17759: gap of unknown length
* 17760 18797: contig of 1038 bp in length
* 18778 18877: gap of unknown length
* 18879 19955: contig of 1078 bp in length
* 19956 20035: gap of unknown length
* 20036 20931: contig of 896 bp in length
* 20932 21011: gap of unknown length
* 21012 21643: contig of 632 bp in length
* 21644 21723: gap of unknown length
* 21724 23416: contig of 1693 bp in length
* 23417 23496: gap of unknown length
* 23497 24457: contig of 961 bp in length
* 24458 24537: gap of unknown length
* 24538 25955: contig of 1418 bp in length
* 25956 26035: gap of unknown length
* 26036 27838: contig of 1803 bp in length
* 27839 27918: gap of unknown length
* 27919 29686: contig of 1768 bp in length
* 29687 29766: gap of unknown length
* 29767 30576: contig of 1210 bp in length
* 30977 31056: gap of unknown length
* 31057 32460: contig of 1404 bp in length
* 32461 32540: gap of unknown length
* 32541 34644: contig of 2104 bp in length
* 34645 34724: gap of unknown length
* 34725 36341: contig of 1617 bp in length
* 36342 36421: gap of unknown length
* 36422 37864: contig of 1443 bp in length
* 37865 37944: gap of unknown length
* 37945 39424: contig of 1480 bp in length
* 39425 39504: gap of unknown length
* 39505 41341: contig of 1837 bp in length
* 41342 41421: gap of unknown length
* 41422 42970: contig of 1549 bp in length
* 42971 43050: gap of unknown length
* 43051 44119: contig of 1369 bp in length
* 44120 44499: gap of unknown length
* 44500 46484: contig of 1985 bp in length

46564: gap of unknown length
48446: contig of 1882 bp in length
48526: gap of unknown length
50593: contig of 2167 bp in length
50773: gap of unknown length
53578: contig of 2805 bp in length
53658: gap of unknown length
55753: contig of 2095 bp in length
55834: gap of unknown length
58457: contig of 2624 bp in length
58537: gap of unknown length
60171: contig of 1634 bp in length
60251: gap of unknown length
63621: contig of 3370 bp in length
63701: gap of unknown length
66219: contig of 2518 bp in length
66299: gap of unknown length
69074: contig of 2775 bp in length
69154: gap of unknown length
72905: contig of 3751 bp in length
72985: gap of unknown length
75963: contig of 2978 bp in length
76043: gap of unknown length
78666: contig of 2823 bp in length
78946: gap of unknown length
82045: contig of 3099 bp in length
82125: gap of unknown length
85135: contig of 3010 bp in length
85215: gap of unknown length
88550: contig of 3335 bp in length
88630: gap of unknown length
92611: contig of 3981 bp in length
92691: gap of unknown length
98709: contig of 6018 bp in length
98789: gap of unknown length
104147: contig of 5358 bp in length
104227: gap of unknown length
111796: contig of 7569 bp in length
111876: gap of unknown length
118488: contig of 6612 bp in length
118568: gap of unknown length
126924: contig of 8356 bp in length
127004: gap of unknown length
137051: contig of 10047 bp in length
137131: gap of unknown length
148859: contig of 11728 bp in length
148939: gap of unknown length
149554: contig of 615 bp in length
149634: gap of unknown length
150397: contig of 563 bp in length
150277: gap of unknown length
150907: contig of 630 bp in length
150987: gap of unknown length
151646: contig of 659 bp in length
151726: gap of unknown length
152074: contig of 348 bp in length
152154: gap of unknown length
152902: contig of 748 bp in length
152982: gap of unknown length
153624: contig of 642 bp in length
153704: gap of unknown length
154242: contig of 538 bp in length
154322: gap of unknown length
154739: contig of 417 bp in length
154819: gap of unknown length
155338: contig of 519 bp in length
155418: gap of unknown length
155993: contig of 575 bp in length
156073: gap of unknown length
156806: contig of 733 bp in length
156886: gap of unknown length
157541: contig of 655 bp in length
157621: gap of unknown length

```

* 157622 158208: contig of 587 bp in length
* 158209 158288: gap of unknown length

alignment_scores:
  Quality: 1257.50      Length: 1432
  Ratio: 1.642          Gaps: 57
  Percent Similarity: 53.492  Percent Identity: 29.190

alignment_block:
US-09-294-298-4 x AC012161/rev ..
Align seg 1/1 to reverse of: AC012161 from: 1 to: 202741

70 PheLeuSerArgArgLeuLysSerSerIleLysArgThrLysSerGlnPr 86
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126785 TTCTTTGGCAGAGA...TCCAACTCTTGAAGCGGACCAAGTCGGTGAC 126739
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

86 oLysLeuAspArgThrSerSerPheArgGlnIleLeuProArgPheArgS 103
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126738 CAAGCTGGAGCGGACCAAG..... 126720

103 erAlaAspHisAspArgAlaArgLeuMetGlnSerPheLysGluSerHis 119
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126719 .....CGCGGATCCGGCGGACTGAGGGGCTCCCGC 126690

120 SerHisGluSerLeuLeuSerProSerSerAlaAlaGluAla..... 133
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126689 TCGCAGGAGAGTCTGCTCCAGTCACGCCGCTCATGTCCACCAATAGGTGA 126640

133 ..... 133

126639 GTTCTGGCTCATATGAGAGTTTGGCCAAACCCCTCAAAACCATTTTCTTTT 126590

134 .....LeuGluLeuAsnLeuAspGluAspSerIleIleLysPro 146
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
12589 TCTTTCTTCTTCTAGATCTCTCGTGCATCGGGCGGTGGCGGCC 126540

147 ValHisSerSerIleLeuGlnGluPheCysPheGluValThrThrse 163
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126539 GTGCATCAGTCGTTCTGGGAGCTGCTCCTCTTCCAGGTACGGGCGG 126490

163 r.....SerGlyThrLysCysPheAlaCysArgSerAlaAlaGluAspL 179
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126489 GCCTCGTGGCGGCGGTACTACTATCGGATCGCGCGGAGCGCGGAC 126440

179 ysrpIleGluAsnLeuGlnArgAlaValLysProAsnLysAspAsnSer 195
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126439 TTTGGATCTACTCGCTGGCAGATCGATCTCCGAATGCGAGACACAG 126390

196 ArgArgValAspAsnValLeuLysLeuTrpIleIleGluAlaArgGluLe 212
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126389 CGTCCGCGGAGCACTCGCTGAAGATGTGGGTGTACGAGGCGAAATCT 126340

212 uProProLysLysArgTyrTyrCysGluLeuCysLeuAspAspMetLeu 229
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126339 CGCGCCCAAGACGCTTACTTTTGGCACTGCAATTTGGACAAGCGTGT 126290

229 yrAlaArgThrThrSerLysProArgSerAlaSerGlyAspThrValPhe 245
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126289 ACGCCGCGACTTCGGTGAAGTGCAGAG.....GATCTGCTGTTT 126249

246 TrpGlyGluHisPheGluPheAsnAsnLeuProAlaValArgAlaLeuAr 262
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126248 TGGGGGAGCACTTTCGATTTCCCGCACATACCGAGATTAATGTATC 126199

262 gLeuHisLeuLysArgSerAspLysArgLysAspLysAlaG 279
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126198 TGTATAGTTTTCGGTGGAGTGGACAGAAAGAGCGGACCAATACC 126149

279 lyPyrValGlyLeuValThrValProValAlaThrLeuAlaGlyArgHis 295
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126148 AATTTTGGGATCGGTGAAGATACCGGTGCACGATGTCCACCTCCAGATTG 126099

```

```

296 PheThrGluGlnTrpTyrProValThrLeuProThrGlySerGlyGlyse 312
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126098 CCCTCCGAACAATGGTATCCATA...CTGACGCAAGCGGCGGACACAG 126052

312 rGlyGlyMetGlySerGlyGlyGlySerGlyGlyGlySerGlyGly 329
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126051 TCTGGGACGACATCGGC.....GGCGGCGGAGTGGGT 126017

329 lyLysGlyGlyGlyGlyCysProAlaValArgLeuLysAlaArgTyrGln 345
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
126016 CCAAGGACAAGAGCAATTCGCCACGCTGAGGATCAAGTGTCTTCCAG 125967

346 ThrMetSerIleLeuProMetGluLeuTyrLysGluPheAlaGluTyrVa 362
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125966 AGCACCCGACATCTGCCCATCANTGTACGCAACTTTTACGCTACCT 125917

362 lThrAsnHisTyrArgMetLeuCysAlaValLeuGluProAlaLeuAsnV 379
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125916 CAAGGAGAACTATAAGCGCGTGTGCGAGACCCCTGGAGCGGTATCGGAG 125867

379 alLysGlyLysGluGluValAlaSerAlaLeuValHisIleLeuGlnSer 395
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125866 TCAAGGCCAAGGAGGACATTTGGACAGCACTGCTGCTGATGCACGCA 125817

396 ThrGlyLysAlaLysAspPheLeuSerAspMetAlaMetSerGluValAs 412
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125816 CAGGATTTGGCGGCGCTTCTCCACCGATGTGGTGGCCCTCGATCTGCT 125767

412 pArgPheMetGluArgGluHisLeuIlePheArgGluAsnThrLeuAla 429
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125766 CGAGTTTGGCGAT...CAGAGGCTTACGTTACGGGGCAACTCTTGGCCA 125720

429 hrLysAlaIleGluLeuTyrMetArgLeuIleGlnLysTyrLeuLys 445
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125719 CCAAGACTGGAGGCAATCTCCTCAGCTGCGGCGGCAACAGTATCTGAG 125670

446 AspAlaIleGlyGluPheIleArgAlaLeuTyrGluSerGluGluAsnCy 462
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125669 GACACATATCCGCACCAATAAAGAGCTAATTCAGTCGGAGAGGAGCTG 125620

462 sGluValAspProIleLys.....CysThrAlaSerSerLeuAlaG 476
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125619 CGAGTGGATCCCAACCAAGACGAGGCTTCTCGCGGGTTCCTGCGAC 125570

476 luHisGlnAlaAsnLeuArgMetCysGluLeuAlaLeuCysLysVal 492
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125569 GACAGCAGCGCGCTTCTCGTGGCGGCTCGCGGCGCTGGCAGTGCATC 125520

493 ValAsnSerHisCysValPheProArgGluLeuLysGluValPheAlaSe 509
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125519 TTCGAATCGCACACAGCAATTTCCCGCCAGTTACGAAATTTGTTTGGAC 125470

509 rTrpArgLeuArgCysAlaGluArgGlyArgGluAspIleAlaAspArgL 526
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125469 GTTCCGGAGCGCTTGCACGACGCTGGCGCTCAGGATATGCGGCAACC 125420

526 euIleSerAlaSerLeuPheLeuArgPheLeuCysProAlaIleMetSer 542
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125419 TGATCTCGGCGAGCAATTTCTCGGCTTCTGCGCCCGGCGCATCTGTCG 125370

543 ProSerLeuPheGlyLeuMetGlnGluTyrProAspGluGlnThrSerAr 559
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125369 CCGTCTCTCTCAATATCACCAGCACTCGCTCGCACGCGGCGGCGGAC 125320

559 gThrLeuThrIleAlaLysValIleGlnAsnLeuAlaAsnPheSerL 576
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
125319 CAATCTCACACTGGTGGCAAGACCTGCAACATTTGGCAATTTTCAACC 125270

576 yspHeThrSerLysGluAspPheLeuGlyPheMetAsnGluPheLeuGlu 592
: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
125269 GCTTCCAGGCAAGAGAACTTTTACAGTTTCTCAACGATTTCTCTCGAG 125220

593 LeuGluTrpGlySerMetGlnGlnPheLeuTyrGluIleSerAsnLeu... 608

```

```

125219 CAGGAGCCGCTCGCATGCAACAGTTCTTGGAGATTATATCCAGCGGCC 125170
609 .....AspThrLeuThrAsnSerSerPheGluGlyT 620
125169 GGAGACCCAGCTCCAGACTCGATCCTCGAT .....TGGCCGGGT 125129
620 yrIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLeuTrpGlu 636
125128 ACATCGACCAAGGCAACAGAGTTGCTCATACACAGTTTGTCTCAGCGAA 125079
637 ValLeuProGlnLeuSerLysGluAlaLeuLeuLysLeuGlyProLeuPr 653
125078 AGCTGCGCAAGCTGCGGAGCCAGGACGACGAGCTGGATCGTTGCA 125029
653 oArgLeuLeuSerAspIleSer .....ThrA 662
125028 GCATATTCTCGATGAATCAGCGGAGCCAAAGAGCATGGCATGGGCACAG 124979
662 laLeu ..... 663
124978 CACTGCGCGGTGGATATTTCGGCGCCACCTCGTCTACGCACCTCGATAGCC 124929
664 .....ArgAsnPro ..... 666
124928 AGCGAGAAATCAGGAGAAATCGAAATCGGGATCCTCGGGCTCGACGCTGG 124879
667 ....AsnIleGlnArgGlnProSerArgGlnSerGluArgAlaArgSerG 682
124878 CTCCAACTCGGAGCAGTTACTGCGCAACAACAAAGCCAGTTGGCCACGCGC 124829
682 lnProMetValLeuArgGlyPro .....SerAlaGlu .....MetGln 694
124828 AGCATGCGGATGTAGTAAACCATATTCTCGCGAGCGGGCATCATGCGA 124779
695 Gly.TyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerPheM 711
124778 GGAGTACTACGCGGAATCTCTGGAGAGAATATCTTTAGATACAATGA 124729
711 etAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerPro 727
124728 TCCACGCGTTAATG ..... 124715
728 ThrLysGluLysProProProProProGlyGlyLysAspLeuPh 744
124715 ..... 124715
744 eTyrValSerArgProProLeuAlaArgSerSerProAlaTyrCysThrS 761
124714 .....GCTTACTCGACGA 124702
761 erSerSerAspIleThrGluProGluGlnLysMetLeuSerValAsnLys 777
124701 GCAGCAGCAGCAGCAGCAGC .....AGCAGCAACAAC 124670
778 SerValSerMetLeuAspLeuGlnGlyAspGlyProGlyGlyArgLeuAs 794
124669 AGCAACAGCAACATC .....AGCAGCTGCAG 124644
794 nSerSerSerValSerAsn .....LeuAlaAlaValGly..AspLeuLe 808
124643 CAGCATGGCCATCAGCAACAGCGCCAGCCAGCATCCAGATCCAGATGCT 124594
808 uHisSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaP 825
124593 CTCCAATTTCACAACTCCCAATT...GCCGGCAACCAATATGAGTTGCTGC 124547
825 roAlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAlaGly 841
124546 CAGGAGGCGCTGCAGCATGCCCAATCGACAGCTCGATGCGTCCCTCATCG 124497
842 MetArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnGl 858

```

```

124496 CTT ..... 124494
858 nLeuArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaA 875
124494 ..... 124494
875 spGlyProGlyProProAlaGlyHisGlyLysSerSerGly .....His 889
124493 .....AATGGGAGCAGCAGCAATTTGCTGTCGCA 124467
890 GlyProProSerSerHisHisHisHisHisHis .....HI 902
124466 GGC .....CACCAGCAGCATGCCATCACCAGCAGCAACTGCA 124429
902 sHisHisArgGlyGlyGluProProGlyAspThrPheAlaProPheHisG 919
124428 TCCACATCAGTCG .....CCGCCG ..... 124410
919 lyTyrSerLysSerGluAspLeuSerThrGlyValProLysProProAla 935
124409 .....GCCGCCACAGCAGCTGCC 124392
936 AlaSerIleLeuHis .....SerHisSerTyrSerAspGl 947
124391 TCCAGCACTATGGAGCGCATGGATGCATCACTATCCGTATATGTGCGA 124342
947 uPheGlyProSerGlyThrAspPheThrArgGlnLeuSerLeuGlnA 964
124341 C .....AATGGCAATGACTACGAG ..... 124323
964 spAsnLeuGlnHisMetLeuSerProProGlnIleThrIleGlyProGln 980
124322 .....ACCAGCAGCCTTCGAGCAGCTCGCTCCAGGACA 124290
981 ArgProAlaProSerGlyPro .....GlyGlyGlySerGlyGlyGlySe 995
124289 CTGCCACGGAATGGAAATCCCAATGCCAATGGCAAGTGGCAGCAGCAA 124240
995 rGlyGlyGlyGlyGlyGlnProProPro.LeuGlnArgGlyLysSer 1011
124239 CAATAACACAGAGCGCAGCTACGATGACATGCACGGGGAGTTCCAANTCC 124190
1012 GlnGlnLeu .....ThrValSerAlaAlaGlnLysProAr 1023
124189 AGATCTCTGGTTCGATACGAGCAGTGTCTTGTCTGCAAGTCGCCACA 124140
1023 gPro .....Ser.GlyAsnLeuLeuLeuGlnSerPro 1033
124139 CCATGATGAATCCAGTTTGGGACAGCGGGCGCGGAGCAAGTCTATCA 124090
1034 GluProSerTyrGlyProAlaArgProArg ..... 1043
124089 CAACCTGAATTTGGGAATACCGCATCACTCAGGTGGCTATGTCCGGGGA 124040
1044 .....GlnGlnSerLeuSerLysGluGlySerIleGlyGlySerGlyG 1058
124039 ATAATTTGAATCCCACTCGAATATGCCAAGAAGTTCGGAGGATCTGGAC 123990
1058 lySer .....GlyGlyGlyGlyGlyGlyLeuLysPro 1069
123989 GATCTGTTCAGTAGCCGAGGAGCAGTGACGTGGCGGAACACAGCAACCA 123940
1070 SerIleThrLysGlnHisSer...GlnThrProSerThr.LeuAsnPro. 1084
123939 TCACAAATCACACACCGGTTCAGCAGACACCCAGGACACTCTGAAGCCGG 123890
1085 ..ThrMetProAlaSerGluArgThrValAlaTrpValSerAsnMetPro 1100
123889 CGCGCGTTCGCGCAAGGAGCAGCTGTGCGGCAAAAGCAGT ..... 123849
1101 HisLeuSerAlaAspIleGluSerAlaHisIleGluArgGluGlnTyrLy 1117
123848 CACTCGAGTTCTGGCTACCGAGCAGTCTCCACAATCCCC ..... 123810

```

```

1117 sLeuLysGluTyrSerLysSerMetAspGluSerArgLeuAspArgValL 1134
      ||| |||:||||:|||||
123809 .....TCGCCCTCGAGTCTCCAGTCCCGTGGAG..... 123780
1134 ysGluTyrGluGluGluIleHisSerLeuLysGluArgLeuHisMetSer 1150
      |||:|||||
123779 .....AGCCAGCTGCAAGGCCGCGATGGCAGTCAC 123750
1151 AsnArgLysLeuGluGluTyrGluArgGluArgLeuSerGlnGluGlu 1167
      ||| |||:||||:|||||
123749 AATGCGCGCTG...GCCCTTCAGAAATCCCTCTATCATCTGCTGACGCCA 123703
1167 nThrSerLysIleLeuMetGlnTyrGlnAlaArgLeuGlnSerGluL 1184
      |||:|||||
123702 AACTGGCTCGTCAGATCATCGCACAGAGTAATACACACAGCAGCAGC 123653
1184 ysArgLeuArgGlnGlnGlnValGluLysAspSerGlnIleLysSerIle 1200
      |||:||||:|||||
123652 AGCAGCAGCAACAACAACAGTTT.....GCCAGCGCGCTTGAACCAATT 123609
1201 IleGlyArgLeuMetLeuValGluGluGluLeuArgArgAspHisProAl 1217
      ||| |||:||||:|||||
123608 GCGCGCGGACTGTCGCGCGAGGCGGCTTTCCTCAACAGTCGCGGAGC 123559
1217 aMet..AlaGluProLeu.....ProGluProLys..... 1226
      |||:||||:|||||
123558 CTTGGAGCGGCGCACATTGACGCCCGAGCTTCGAGCGAAGCAGCTCTCGG 123509
1227 .....LysArgLeuLeuAspAlaGlnArgGlyLe 1236
      ||| |||:||||:|||||
123508 CGGATAATTACTTCAGTTATGTCAGCGGCTGCGAGCTGCTGGAGCGGTATT 123459
1236 rPheProTrpValGlnGlnThrArgValArgTrpProHisLeuGlyT 1253
      ||| |||:||||:|||||
123458 GCACCAAAATTGGAGCTCAACGCTCGC.....TCAGCGCGCGGAG 123418
1253 hrAlaTrpProGln..... 1258
      ||| |||:||||:|||||
123417 TAGCTCTCCACCTCAGCATCTGCTGCCACCTCGAATCTTGCAAGAGCG 123368
1259 ....ProHisProHisProGlyCysArgSerGlnArgThrAlaSerSe 1274
      ||| |||:||||:|||||
123367 GCGCTCATCGGCTACGCGCGGCTGATGGCCGCTGCAAGCGCAGGAT 123318
1274 rGlyThrProGlnThrSerProPro.....SerIleThrA 1287
      ||| |||:||||:|||||
123317 GTCTAGCGCAGTGGCTACGCGCGCAGCAGTGGAAATGTGGCTATGGCTT 123268
1287 sLeuLeuPro..Cys.....AlaProTyrProGlyProGlyProSerValTh 1301
      ||| |||:||||:|||||
123267 GTCCATTTCAGTGGCGGGACACATCAATCCACCCAGCAGCAGC 123218
1301 rAspLeuLeuProSerAlaProAspProGlyThrSerProThrThrArgT 1318
      ||| |||:||||:|||||
123217 AGAATCCGATGTCAGCAGCAGCAGCAAGGCAAGGATCAGGACACACAAG 123168
1318 hr.....GlyArgHisGlnGlyThrAlaGlyLeuSerPro 1329
      ||| |||:||||:|||||
123167 CAGTATCGCGCAGTGTGGCGGCGAGCGTGGGATCGGCACATCAGCGGC 123118
1330 LeuThrProPro.....TrpGlyThrHis 1337
      ||| |||:||||:|||||
123117 TCAGAGCGCGCTGAGCTTGGACTCGGCGCGCAC 123085
seq_name: gb_in1:AB011280
seq_documentation_block:
LOCUS AB011280 3984 bp mRNA 25-FEB-1998
DEFINITION Caenorhabditis elegans mRNA for GAP-2-4, complete cds.
ACCESSION AB011280
VERSION AB011280.1 GI:2911175
KEYWORDS .GAP-2-4.
SOURCE
  Caenorhabditis elegans (strain:Bristol N2) mixed stage
  hermaphrodite cDNA to mRNA.
ORGANISM
  Caenorhabditis elegans
  Eukaryota; Metazoa; Nematoda; Secernentea; Rhabditia; Rhabditida;
  Rhabditina; Rhabditoidae; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
  1 (bases 1 to 3984)
  Yamamoto,M., Iino,Y. and Hayashizaki,S.
  Direct Submission
  Submitted (18-FEB-1998) to the DDBJ/EMBL/GenBank databases
  Masayuki Yamamoto, University of Tokyo, Department of Biophysics
  and Biochemistry; 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
  (E-mail:myamamoto@ims.u-tokyo.ac.jp, Tel:81-3-3812-2111,
  Fax:81-3-5802-2042)
  2 (sites)
  Hayashizaki,S., Iino,Y. and Yamamoto,M.
  Characterization of the C. elegans gap-2 gene encoding a novel
  Ras-GTPase activating protein and its possible role in larval
  development
  Genes Cells (1998) In press
FEATURES
  Location/Qualifiers
    1..3984
      /organism="Caenorhabditis elegans"
      /strain="Bristol N2"
      /db_xref="taxon:6239"
      /dev_stage="mixed stage"
      /sex="hermaphrodite"
    20..3382
      /gene="gap-2"
      /gene="gap-2"
      /codon_start=1
      /product="GAP-2-4"
      /protein_id="BAA24960.1"
      /db_xref="GI:2911176"
      /translation="MHTSPPIQEQFOQYVQFQYVSSRHVIFAQSQRRPQRDQRENASI
      RKLRSCHDLSHDHNFSPPIRKKTKHVPSTLSLVTHDAIHDETNSGGAQVGES
      LDRGWSGASTGNRTTLTASVHNHLMGRSSSHNLSTLSTGSLNLOPTNAYGFP
      TVPPANFSDSASGRITANLLSRPRGSLNPLKRTKSVSKMEKSLAEANQSLHVRDASN
      TPRSLSLYAQPAPRHLSOPAREGSLRACRSHESLSASHSLHSHLNDRLHPVH
      PSIFEPNCFRLASTYSCRTPLERAKWMENLRKTMPPRQQRRTENSMIWILEAK
      GLPAKRYVCMTLDKTYAKTSSKARTDNVFGENFEFMLPKIDECVSLPRESDS
      KKKDGLGYVTIGIDOLSSPSRWYNTVNTSHSDSTSRISALGGKSSSOEPLSL
      RIKARQSHVILPLRAYDNLETLCTYNYLPLCEQLPVLNVKDEDLATSLVRYVKH
      NLAKETLCLIMKEVKLNDHLMFRNTLAKMESFMKLADVADLDLSTLSDFIKTY
      LQCEDEQVDFQKLGNSVSLKRNALLMYVEAVTKILNHNHOLPKNLRDVFSL
      RCLRLEAQRREALDTLTSSTIFLFLCPAILSPSLNLSVSEPSPTNARNLTIKATL
      QNLANSKFGGKEPHEMFEFVDREWHMKDFLLRISSEKSGPEKNADAIYDAKE
      LSLIATYLEAWTPLLQKNGKHPLSNVKSLSELAECKRSDNGVFHSPMVOQPS
      DYENSPQOHVVRHENVPAYRSTPTGQATVGLGRSTNRPATLHLLTSDVYVLSAFOTP
      SLRPGTRLSDETGCTSSRTSDKTTSSAEIRDDTDSDFELREDRGGGRNRKRLPTD
      ASPSSQQAASSGYLSNPNRSYSNSSSPVPERMAALSTANFCFGGPGSSGYPAE
      PKRVIQKRASPPPPYDPDVHNYHYPQMVAVPDCVSPRTQRTQGVVNAQNLSLPR
      TNPASRNSTLLRPSVNVNPDWDRTSDYMDRGENNRYRSQLESQVESQARELERLWR
      ENIELKSKRCLOQKLWIPSDLTIVLVRIPTIL"
    polyA_site
      /note="15 A nucleotides"
  BASE COUNT 1167 a 943 c 848 g 1025 t 1 others
  ORIGIN
  alignment_scores:
    Quality: 1172.00 Length: 887
    Ratio: 2.147 Gaps: 26
    Percent Similarity: 61.556 Percent Identity: 34.386
  alignment_block:
    US-09-294-298-4 x AB011280
  Align seg 1/1 to: AB011280 from: 1 to: 3984
  71 LeuSerArgLeuLysSerSer...IleLysArgThrLysSerGlnPr 86
  ||||| |||:||||:|||||
  560 TTATCAGCAGCGTTCCGTTCAACCCCAAGAGAGAGTGTGAG 609

```



```

::: |||
2353 ATGCAAGAGACGCTCGGATAATGCGCTCTTCACTCTCCAATGGTTCAA. 2401
670 tgGlnProSerArgGlnSerGluArgAlaArgSerGlnProMetValLeu 686
||||| |||
2402 ..CAGCGCTCTCAGATTGAGTAATGTCACAGCAACACAGCTGTCTCT 2449
687 Arg.....GlyProSerAlaGluMet 694
||| |||
2450 CCACATGAGATGATACCGGATATCGCAGTACTCCGCCCACTGGCCAAAGC 2499
694 nGlyTyrMetMetArgAspLeuAsnSerSerIleAspLeuGlnSerPheM 711
::: |||::: |||
2500 CACGGTATGGCGCTTCTACAAACCGCCT..... 2530
711 etAlaArgGlyLeuAsnSerSerMetAspMetAlaArgLeuProSerPro 727
||| ||| ::: ||| |||
2531 ..GCACACTCATTTGCTCACATCGGAGCATTAACGTC.....CTCTCT 2569
728 ThrLysGluLysProProProProProGlyGlyGlyGlyAspLeuPh 744
::: ||| ||| ||| |||
2570 TCTGCATTTCAAACTCCAAGTCTTCGCCCTGGAGGC..... 2605
744 eTyrrValSerArgProLeuAlaArgSerSerProAlaTyrCysThrS 761
2606 .....ACTCGTTAACGGATGAACCTGGCACCT 2633
761 erSerSerAspIleThrGluProGluGlnLysMetLeuSerValAsnLys 777
||||| ||| ::: |||
2634 CTTGAGTCGACGACGAGTGAAGAACAACACGAGCGCTGAGATTCGAGAC 2683
778 SerVal...SerMetLeuAspLeuGlnGlyAsp...GlyProGlyGlyArg 792
||| ||| ||| ||| ||| |||
2684 GACACTGATTCGGATTTCGAGTACGAGAGGATCGAGGACGCGGAGGAG 2733
792 gluAsnSerSerSerValSerAsnLeuAlaAlaValGlyAspLeuLeuH 809
| ||| ||| |||
2734 A...AACGTAAGAGGCTACCGACTACTGATGCA..... 2764
809 isSerSerGlnAlaSerLeuThrAlaAlaLeuGlyLeuArgProAlaPro 825
||| ||| ||| |||
2765 ..TCACCATCGACGAGTCAACAGCTTCAAGTGGATATTAAGTAATAAT 2812
826 AlaGlyArgLeuSerGlnGlySerGlySerSerIleThrAlaAlaGlyMe 842
||| ||| ||| ||| ||| |||
2813 CTTGAGATCCACTACTCAACTCTTCGAGTTCATCT..... 2851
842 tArgLeuSerGlnMetGlyValThrThrAspGlyValProAlaGlnGlnL 859
||| ||| ||| ||| |||
2852 .....CCAGTTGAACGAA 2864
859 euArgIleProLeuSerPheGlnAsnProLeuPheHisMetAlaAlaasp 875
::: ||| ||| ||| ||| |||
2865 TG...CGCGCTCATCAATGCTTAACCCATGCTT..... 2896
876 GlyProGlyProProAlaGlyHisGly..... 884
||||| ||| ||| ||| |||
2897 GCACGAGCCCATCATCTGCGATATGCTATACCTGCGAGACCCAAAGGAAT 2946
885 .....GlySerSerGlyHisGlyProProSerSerHisHisHisH 899
::: ||| ||| ||| |||
2947 CGTATACCAAGGAGCAAGCCGACCAACCATACGATCCGATGTCACACA 2996
899 isHisHisHis 902
::: ||| |||
2997 ACTATCATAT 3007

```

seq_name: gb_pr4-AK024488

seq_documentation_block:

LOCUS AK024488 4287 bp mRNA PRI 29-SEP-2000
 DEFINITION Homo sapiens mRNA for FLJ00087 protein, partial cds.
 ACCESSION AK024488

```

VERSION AK024488.1 GI:10440489
KEYWORDS tis (full insert sequence)
SOURCE Homo sapiens adult spleen cDNA to mRNA, clone:as00087.
ORGANISM Homo sapiens
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Eukaryota; Euthera; Primates; Catarrhini; Hominidae; Homo.
TITLE The nucleotide sequence of a long cDNA clone isolated from human
        spleen
JOURNAL Published Only in DataBase (2000) In press
REFERENCE 2 (bases 1 to 4287)
AUTHORS Ohara,O., Nagase,T., Kikuno,R. and Okumura,K.
TITLE Direct Submission
JOURNAL Submitted (24-AUG-2000) to the DDBJ/EMBL/GenBank databases. Osamu
        Ohara, Kazusa DNA Research Institute, Department of Human Gene
        Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan
        (E-mail:cdna@info.kazusa.or.jp, URL:http://www.kazusa.or.jp/NEDO,
        Tel:81-438-52-3913, Fax:81-438-52-3914)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
        International Trade and Industry of Japan; cDNA full insert
        sequencing; Research Association for Biotechnology; cDNA library
        construction, 5'- & 3'-end and one pass sequencing and clone selection:
        Kazusa DNA Research Institute.
FEATURES
        Location/Qualifiers
            1..4287
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="as00087"
                /dev_stage="adult"
                /tissue_type="spleen"
                /notes="vector:pBluescriptII SK plus"
            1339..3363
                /gene="FLJ000087"
                <1339..3363
                /gene="FLJ000087"
                /notes="For this clone, Genemark analysis triggered an
                alert for spurious CDS split. The result of Genemark
                analysis of the nucleotide sequence of this clone is
                accessible through http://www.kazusa.or.jp/NEDO.
                Start codon is not identified."
                /codon_start=1
                /product="FLJ00087 protein"
                /protein_id="BAI15778.1"
                /db_xref="GI:10440490"
                /translation="GSTLHLRCLLWANKRPPSPSHSHLDQNVREETWLSVWVHEA
                KGLPRAAGAPGVRAELWDLGALLARTAPRAGQGLFWAEFHFELPPARRLSRLR
                GLPGSAVLGRVALAELDAPRAPAGLERWFFLLGAPAGAAALRARRRRLVLP
                ERYKELAEPLTFHYARLCALEPALPAQKEELAAAMRVRLATGAAQALVDTGTAE
                LARCGREALFRENTLATKIDYMKLVADYQLETGLQGVRRLCATGECEDVDSK
                CPASLPEHOARLNRSCBEVFETIIHSYDNFPAELTIIAKVIONLANRPPFGEVPM
                VCASLFLRLCLPAILAPSLFGLPADHPAPGAPARTLTIAKVIONLANRPPFGEVPM
                GPMNSFLEHPAMOCFLDOVAMVDVDAAPSCYSGDGLALQALVLAHLCITFAELD
                OTTRDTLEPLTILRAIEGQPLVSVPMRLPLPAQVHSLNAGEPKGFLAERDLPK
                HTPLISKQSLSVRSRSESWARPRDEERLRRPRPQRTQCSVPVRRPARRQSAGFW
                PRPKGSLMGAPAPRATWRDSASLPKPSVPMQRMQDQDRRQALGTHRPVKNLAE
                LQCEAALREEQKVLRSLSLSLTQIRALTEQEQQLRGLQDLDSRLRAG"
BASE COUNT 785 a 1529 c 1261 g 712 t
ORIGIN

```

alignment_scores:

Quality: 1103.50 Length: 1425
 Ratio: 1.695 Gaps: 48
 Percent Similarity: 45.684 Percent Identity: 27.930

alignment_block:

US-09-294-298-4 x AK024488 ..

Align seg 1/1 to: AK024488 from: 1 to: 4287

2 ProProArgProArgGlnGlyAlaProValcInValProCysProLe 18
 ||| |||:::||||::: ||| ||| |||

9

OM of: US-09-294-298-4 to: EST.* out_format : pfs

Date: Jan 18, 2001 9:24 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL-frame.p2n.model -DEV-xlp
-Q/cgn2/USPTO.spool/US09294298/runat_17012001_125501_27745/app_query.fasta.1.4359
-DB-EST -QFMT=fastap -SUFFIX=1st -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.500 -CGAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-GAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blossum62 -TRANS-human4.0.cdf
-LIST=45 -DOALIGN=200 -THR_SCORE=PCT -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -MINLEN=0
-MAXLEN=2000000000 -USER=US09294298 -CGN1_1/119 -NCPU=6
-ICPU=3 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-294-298-4

Query length: 1376

Database: EST.*

Database sequences: 7991742

Database length: -791223438

Search time (sec): 4688.750000

score_list:

| Sequence | Strd | Orig | Zscore | BScore | Len | Documentation |
|--------------------|------|--------|---------|---------|------|----------------------------------|
| gb_est37:AV590698 | + | 933.00 | 1064.14 | 2.9e-50 | 571 | AV590698 Bos taurus b2 |
| gb_gss21:AX399131 | - | 764.00 | 872.80 | 1.3e-39 | 474 | AX399131 NM016414R Mouse 10XB |
| gb_est49:AW779747 | - | 737.50 | 839.91 | 9.0e-38 | 593 | AW779747 hm85f10.x1 NCI_CGAP_K1 |
| gb_est44:AW421227 | - | 731.50 | 832.96 | 2.2e-37 | 597 | AW421227 fj94f10.x1 zebrafish g |
| gb_est71:BB490923 | - | 706.00 | 803.20 | 1.0e-35 | 628 | BB490923 db38g08.x1 Blackshear/ |
| gb_est22:AI210671 | - | 660.00 | 751.72 | 7.4e-33 | 566 | AI210671 ml58d01.y1 Stratagene |
| gb_est71:BB463433 | - | 628.50 | 717.95 | 5.6e-31 | 463 | BE463433 hw23d06.x1 NCI_CGAP_K1 |
| gb_est22:AI210671 | - | 600.00 | 684.20 | 4.2e-29 | 511 | AI210671 UT-R-YO-vk-d-04-0-UT-s |
| gb_est22:AI210671 | - | 571.50 | 652.43 | 2.5e-27 | 474 | AI205989 UT-H-BII-afy-d-12-0-UT |
| gb_est28:AS396549 | + | 566.50 | 642.16 | 9.3e-27 | 707 | AS396549 AJ396547 dmz426 Gallu |
| gb_est11:AA709010 | - | 563.00 | 643.89 | 7.5e-27 | 427 | AA709010 z94h11.s1 Soares_pine |
| gb_est22:AI210671 | - | 548.50 | 626.44 | 7.0e-26 | 460 | AI210671 t91609.x1 NCI_CGAP_K1 |
| gb_est9:AA554056 | - | 543.50 | 621.00 | 1.4e-25 | 449 | AA554056 ml01g07.s1 NCI_CGAP_K1 |
| gb_est27:AA159829 | - | 537.50 | 611.71 | 4.6e-25 | 555 | AI210671 f22g01.y1 zebrafish w |
| gb_est40:AW136165 | - | 531.50 | 607.41 | 8.0e-24 | 452 | AW136165 UT-H-BII-abz-d-09-0-UT |
| gb_est9:AA554055 | - | 517.50 | 597.16 | 6.5e-24 | 452 | AA554055 ml01g06.s1 NCI_CGAP_K1 |
| gb_est10:AA691827 | + | 504.00 | 576.73 | 4.1e-23 | 413 | AA691827 vt05d01.r1 Barstead md |
| gb_gss26:CN502VGV | + | 501.00 | 563.00 | 2.4e-22 | 1020 | AL215753 Tetradodon nigroviridis |
| gb_est74:BB695645 | + | 486.00 | 553.59 | 8.0e-22 | 516 | BB695645 MRI-BR0800-030700-0021 |
| gb_est9:AA575948 | - | 484.50 | 555.39 | 6.3e-22 | 379 | AA575948 nm56d05.s1 NCI_CGAP_K1 |
| gb_est22:AI210671 | - | 484.50 | 555.30 | 6.4e-22 | 382 | AI210671 vt73f04.x1 Stratagene |
| gb_est41:AW204612 | + | 481.00 | 552.38 | 9.3e-22 | 347 | AW204612 UT-H-BII-aej-c-02-0-UT |
| gb_est41:AA107246 | + | 480.00 | 548.54 | 1.5e-21 | 440 | AA107246 ml58d01.r1 Stratagene |
| gb_est39:AW076911 | + | 473.00 | 543.29 | 3.0e-21 | 345 | AW076911 vx39f05.y1 Soares.thym |
| gb_est76:BB848541 | + | 464.00 | 526.76 | 2.5e-20 | 596 | BB848541 ut39f05.y1 Soares.thym |
| gb_gss26:CN5023DV | + | 461.50 | 518.25 | 7.4e-20 | 979 | AL200092 Tetradodon nigroviridis |
| gb_gss26:CN5023KG | + | 457.50 | 515.52 | 1.1e-19 | 832 | AL200092 Tetradodon nigroviridis |
| gb_gss14:AA099396 | + | 454.00 | 513.78 | 1.3e-19 | 682 | AQ993964 RFI-23-24C10-TV RPTC |
| gb_gss26:CN501HQ5 | + | 452.50 | 512.95 | 1.5e-19 | 631 | AL172022 Tetradodon nigroviridis |
| gb_gss26:CN501HWF | - | 425.50 | 476.73 | 1.5e-17 | 1006 | AL073826 Drosophila melanogast |
| gb_gss26:CN503ANH | + | 419.50 | 470.22 | 3.5e-17 | 975 | AL235430 Tetradodon nigroviridis |
| gb_est26:AA527364 | - | 418.50 | 481.87 | 7.9e-18 | 317 | AA527364 UT-R-B01-ajj-e-01-0-UT |
| gb_est59:BB321064 | + | 417.50 | 479.18 | 1.1e-17 | 363 | BB321064 BB321064 RIKEN full-le |
| gb_est26:AA1874961 | + | 416.50 | 478.92 | 1.2e-17 | 336 | AI874961 ul27f03.x1 Sugano.mous |
| gb_gss27:AA504DL0 | + | 416.50 | 467.99 | 4.7e-17 | 877 | AL285865 Tetradodon nigroviridis |
| gb_gss27:AA504DL0 | + | 401.50 | 462.34 | 9.7e-17 | 319 | AA258366 z956d01.s1 Soares.Nhm |
| gb_gss27:CN504LW | + | 401.00 | 449.63 | 4.9e-16 | 926 | AI270761 Tetradodon nigroviridis |
| gb_est47:AA656799 | + | 394.00 | 447.71 | 6.3e-16 | 542 | AA656799 109245 MARC 180V Bos t |
| gb_est61:BB394547 | + | 383.50 | 441.52 | 1.4e-15 | 325 | BB394547 BB394547 RIKEN full-le |
| gb_est23:AI650331 | + | 379.00 | 437.43 | 2.4e-15 | 296 | AI650331 wal8f01.x1 NCI_CGAP_K1 |
| gb_est14:AA953060 | - | 368.50 | 425.57 | 1.1e-14 | 292 | AA953060 co01g09.s1 Soares.NFL |

gb_est59:BB312776 + 365.50 421.49 1.8e-14 309
gb_est75:BE773015 - 365.50 416.45 3.5e-14 481
gb_est19:AI327335 - 358.00 404.42 1.6e-13 651
gb_est66:BB085114 - 357.50 411.41 6.6e-14 335

seq_name: gb_est37:AV590698

seq_documentation_block:

LOCUS AV590698 571 bp mRNA EST 29-AUG-2000
DEFINITION AV590698 Bos taurus brain fetus Bos taurus cDNA clone EIBR014B04
5', mRNA sequence.

ACCESSION AV590698

VERSION AV590698.1 GI:9701691

KEYWORDS EST.

SOURCE cow.

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 571)

AUTHORS Sugimoto,Y., Hirotsune,S., Takasuga,A., Itoh,R., Jitohzono,A. and
Suzuki,H.

TITLE bovine cDNA sequencing

JOURNAL unpublished (2000)

COMMENT Contact: Yoshikazu Sugimoto

Animal Genetics Division

Shirakawa Institute of Animal Genetics

Odakura, Nishigo, Nishi-shirakawa, Fukushima 961-8u1, Japan

Tel: 81-248-25-5641

Fax: 81-248-25-5725

Email: kazusugi@cococ.ocn.ne.jp

Single pass sequencing.

This clone was obtained from a polyA-deleted cDNA library.

FEATURES

source location/Qualifiers

1..571

/organism="Bos taurus"

/db_xref="taxon:9913"

/clone="EIBR014B04"

/tissue_type="brain"

/dev_stage="fetus"

/lab_host="DH108"

/note="vector: pZL1; Site:1: SalI; Site:2: NotI; Poly: A

was deleted from a NotI site"

BASE COUNT 124 a 161 c 169 g 116 t

ORIGIN 1 others

alignment_scores:

Quality: 933.00 Length: 186

Ratio: 5.043 Gaps: 1

Percent Similarity: 99.462 Percent Identity: 99.462

alignment_block:

US-09-294-298-4 x AV590698

Align seg 1/1 to: AV590698 from: 1 to: 571

382 LyscluvAlA1aSerAlaLeuValHisIleLeuGlnSerThrGlyLy 398

|||||

13 AAGGAGAGTCCGCCAGTCTCTGGTTCACATCCGAGAGTACAGCAA 62

|||||

398 sAlaLysAspPheLeuSerAspMetAlaMetSerGluValAspArgPheM 415

|||||

63 GGCCAGGAGCTCTCTTTCAGACATGCCCATGTCGAGGTGGACCTTCA 112

|||||

415 etGluArgGluHisLeuIlePheArgGluAsnThrLeuAlaThr 441

|||||

113 TGAACGAGACACACCTCATATTCCCGAGAACACGCTCCACATAAA 162

|||||

432 ileGluGlyTyrMetArgLeuIleGlyGlnTyrLeuTyrAspAlaIle 448

|||||

163 ATAGAGAGTACATGAGACTGTTTGGTCAGAAATACCTCAGAGTGCAT 212

| Sequence | Strd | Only | Zscore | EScore | Len | ! Documentation |
|--|------|------|--------|--------|--------|-----------------|
| /SID56/cgcgdata/geneseq/geneseq/NA1992.DAT:Q26078 + | | | | 437.50 | 304.80 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1991.DAT:Q10988 + | | | | 435.50 | 303.40 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:Q26739 + | | | | 429.50 | 302.84 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1998.DAT:V62176 + | | | | 411.50 | 192.36 | |
| /SID56/cgcgdata/geneseq/geneseq/NA2000.DAT:Z87283 + | | | | 298.50 | 198.15 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:Z58159 + | | | | 293.50 | 194.67 | |
| /SID56/cgcgdata/geneseq/geneseq/NA2000.DAT:T268980 + | | | | 286.50 | 194.56 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1998.DAT:Z23430 + | | | | 283.00 | 197.93 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1998.DAT:Z62154 + | | | | 274.50 | 179.32 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1995.DAT:Q86297 + | | | | 273.00 | 189.03 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1995.DAT:Q75977 + | | | | 273.00 | 187.42 | |
| /SID56/cgcgdata/geneseq/geneseq/NA2000.DAT:Z87318 + | | | | 270.00 | 172.03 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1996.DAT:T06769 + | | | | 263.50 | 169.37 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T89956 + | | | | 263.50 | 165.31 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1998.DAT:V05287 + | | | | 260.50 | 180.86 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1999.DAT:X56436 + | | | | 259.50 | 163.50 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T00413 + | | | | 253.50 | 159.12 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T78508 + | | | | 253.50 | 159.12 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T00414 + | | | | 253.00 | 174.33 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1992.DAT:Q29263 + | | | | 252.00 | 176.68 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1999.DAT:T06978 + | | | | 252.00 | 176.68 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1999.DAT:X21373 + | | | | 251.50 | 174.82 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1999.DAT:V81748 + | | | | 251.00 | 185.47 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1995.DAT:Q80016 + | | | | 251.00 | 170.19 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1999.DAT:X36583 + | | | | 251.00 | 169.29 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:Q20602 + | | | | 251.00 | 167.95 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T46941 + | | | | 251.00 | 165.96 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T80415 + | | | | 250.50 | 173.48 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1999.DAT:X58463 + | | | | 250.00 | 176.36 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T86704 + | | | | 249.00 | 175.49 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1995.DAT:Q96261 + | | | | 248.50 | 171.31 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1994.DAT:Q84201 + | | | | 247.50 | 173.07 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1997.DAT:T89345 + | | | | 247.50 | 167.86 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1995.DAT:Q84660 + | | | | 247.50 | 167.86 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1995.DAT:Q84659 + | | | | 247.50 | 167.86 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1994.DAT:Q73500 + | | | | 247.50 | 167.28 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1994.DAT:Q73587 + | | | | 247.50 | 167.28 | |
| /SID56/cgcgdata/geneseq/geneseq/NA2000.DAT:Z55881 + | | | | 246.50 | 151.69 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1991.DAT:Q13669 + | | | | 246.50 | 174.38 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1992.DAT:V24130 + | | | | 246.50 | 158.12 | |
| /SID56/cgcgdata/geneseq/geneseq/NA1998.DAT:Q24828 + | | | | 245.50 | 165.74 | |

147 ValHisSerSerIleLeuGlyG

```

alignment_scores:
    Quality: 437.50      Length: 568
    Ratio: 1.326        Gaps: 22
    Percent Similarity: 58.099    Percent Identity: 26.232

alignment_block:
  US-09-294-298~4 x Q26078 ..

Align seg 1/1 to: Q26078 from: 1 to: 4307

147 VAIHSSSRILEUCILGGLGLPHCYSPHECLVALHRRTHRS
|||||:|||||:|||||:|||||:|||||:|||||:

```

```
seq_name: /SID56/gcgdata/geneseq/geneseq/WAL191.DAT:Q10988
seq_documentation_block:
ID: Q10988 standard; cDNA; 4307 BP.
XX
AC: Q10988;
XX
DT: 22-MAY-1991 (first entry)
```


OM of: US-09-294-298-4 to: Issued_Patents_NA:* out_format : pfs
Date: Jan 19, 2001 7:48 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters: -DRV-xlp
-MODE=frame+pz2n.model -US09294298/runat_17012001_125501_27768/app_query.fasta_1.4359
-Q/cgn2_1/US09294298/US09294298 -SUFFIX=ini -GAPOP=12.000
-DB=Issued_Patents_NA -OFMT=fastap -SUFFIX=ini -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -GAPOP=10.000 -XGAPEXT=0.500
-GAPOP=6.000 -GAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cgi -LIST=45 -DOCLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -MINLEN=0 -MAXLEN=200000000
-USER=US09294298 -CGMA_1_163 -NCPU=6 -ICPU=3 -LONGLOG -NO_XLPXY
-WAIT -THREADS=1

Search information block:
Query: US-09-294-298-4
Query length: 1376
Database: Issued_Patents_NA:*
Database sequences: 280836
Database length: 80580151
Search time (sec): 156.930000

score.list:

| Sequence | Strd Orig | ZScore | Escore | Len | Documentation |
|---|-----------|--------|--------|---------|---------------|
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-24 | + | 437.50 | 421.08 | 6,7e-16 | 3456 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-7 | + | 437.50 | 418.91 | 8,8e-16 | 4307 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-909-954-3 | + | 371.00 | 355.38 | 3,0e-12 | 3117 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-909-954-1 | + | 362.50 | 346.76 | 9,2e-12 | 3148 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-557-139-1 | + | 283.00 | 260.40 | 5,9e-07 | 6152 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-258-261B-6 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-456-837-6 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-342-6 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-457-646A-6 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-458-076A-6 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-4 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-729-211-6 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-09-028-934-6 | + | 263.50 | 225.58 | 5,2e-05 | 28958 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-764-233A-1 | + | 263.50 | 220.32 | 0.0001 | 49377 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-1 | - | 259.50 | 217.61 | 0.0001 | 43280 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-7 | - | 253.50 | 211.34 | 0.0003 | 44377 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-198-1 | - | 253.50 | 211.34 | 0.0003 | 44377 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-781-802-7 | - | 253.00 | 236.91 | 1,2e-05 | 3147 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-551-531-1 | + | 251.00 | 228.45 | 3,6e-05 | 6060 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-449-933-1 | + | 251.00 | 224.62 | 5,8e-05 | 8937 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-411-389-1 | + | 251.00 | 222.84 | 7,3e-05 | 10706 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-574-959A-6 | + | 248.50 | 230.28 | 2,8e-05 | 3901 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-13 | - | 248.50 | 227.09 | 4,3e-05 | 5392 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-510-646B-1 | + | 248.50 | 227.09 | 4,3e-05 | 5392 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-949-386-23 | + | 247.50 | 222.46 | 7,7e-05 | 7791 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-450-562-23 | + | 247.50 | 222.46 | 7,7e-05 | 7791 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-949-386-22 | + | 247.50 | 222.44 | 7,7e-05 | 7808 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-450-562-22 | + | 247.50 | 222.44 | 7,7e-05 | 7808 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-07-945-283-1 | - | 247.50 | 221.67 | 8,5e-05 | 7808 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-09-335-409-1 | + | 247.50 | 201.01 | 0.0012 | 68750 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-149-097D-22 | + | 247.00 | 221.94 | 8,2e-05 | 7808 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-227C-7 | + | 245.00 | 202.81 | 0.0010 | 44377 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-804-198-1 | + | 245.00 | 202.81 | 0.0010 | 44377 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-779-801-3 | + | 243.50 | 225.96 | 4,9e-05 | 3632 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-343-785-7 | + | 242.50 | 230.59 | 2,7e-05 | 2051 |
| /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-462-221-7 | + | 242.50 | 230.59 | 2,7e-05 | 2051 |
| /cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-946-458-7 | + | 242.50 | 230.59 | 2,7e-05 | 2051 |
| /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-343-785-1 | + | 242.50 | 230.31 | 2,8e-05 | 2111 |

/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-462-221-1 + 242.50 230.31 2,8e-05 2111
/cgn2_6/ptodata/2/ina/6_COMB.seq:US-08-946-458-1 + 242.50 230.31 2,8e-05 2111
/cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-574-959A-8 + 242.00 225.67 5,1e-05 3211
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-14418-1 + 241.50 225.69 5,1e-05
/cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:PCT-US95-15327-1 + 241.50 225.69 5,1e-05
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-190-687B-24
seq_documentation_block:
; Sequence 24, Application US/08190687B
; Patent No. 5760203
; GENERAL INFORMATION:
; APPLICANT: Wony, Gail L.
; APPLICANT: Martin, George
; APPLICANT: McCormick, Francis P.
; APPLICANT: Rubinfeld, Bonnie
; APPLICANT: O'Rourke, Edward C.
; APPLICANT: Clark, Robin
; TITLE OF INVENTION: GAP Gene Sequences
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60608-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,687B
; FILING DATE: 02-FEB-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/774,644
; FILING DATE: 11-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/260,807
; FILING DATE: 21-OCT-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/230,761
; FILING DATE: 10-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 27527/31898
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3456 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 100..2709
; US-08-190-687B-24

alignment_scores:
Quality: 437.50 Length: 568
Ratio: 1.326 Gaps: 22
Percent Similarity: 58.099 Percent Identity: 26.232

alignment_block:
US-09-294-298-4 x US-08-190-687B-24

Align seg 1/1 to: US-08-190-687B-24 from: 1 to: 3456

```

147 ValHisSerIleLeuGlyGlnGluPheCysPheGluValThrThrse 163
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1156 GTTCATGATAGTCTCTTTGGAGGCGCAAACTGTTTCAGATAGTATCA 1205
163 rSerGlyThrLys.....CysPheAlaCysArgSerAlaA 175
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1206 GCACATTAGTAGAACACATTACATCTTTTACTTTGCAGGAGAACTCCAG 1255
175 laGluArgAspLysTrpIleGluAsnLeuGln.....Arg 186
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
1256 AACAGCAGCAGGATTGGATGAAGAGTTCGAGCGCATTTTCAATTACGG 1305
187 AlaValLysProAsnLysAspAspSerArg...ArgValAspAsnValLe 202
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1306 AAAAGTAGTCCAGGACATCCAAATAAAGCGCTTCGTCAGGTCAGAGCCT 1355
202 uLysLeuTrpIleIleGluAlaArgGluLeuProProLys.....LysA 217
| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1356 TGTTTTACATATTGAAGAGCGCCATAAATCCAGATAAAACATTTTACTA 1405
217 rGlyTrpCysGluLeuCysLeuAspAspMetLeuTrpAlaArgThrThr 233
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1406 ATCCATATTGTAACATCTACCTGAATAGTGCCAAAGTAGCAAACTCAT 1455
234 SerLysProArgSerAlaSerGlyAspThrValPheTrpGlyGluHisph 250
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1456 GCAAGG.....GAAGGCAAAACCCAGTATGGTCAGAAAGATT 1493
250 eGluPheAsnAsnLeuProAla...ValArgAlaLeuArgLeuHisLeuT 266
| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1494 TGTCTTTGATGATCTTCCTCTGACATCAATAGATTGAAATCACTCTT. 1542
266 yrArgAspSerAspLysLysArgLysAspLysAlaGlyTrpValGly 282
: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1543 .....AGTAATAAACAAGAAAGCAAAAGATCTGATATCTTA 1581
283 LeuValThrValProValAlaThrLeuAlaGlyArgHisPheThrGluG1 299
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1582 TTTATGCGTGCGCATGAGCGGATTACAGAAAGGCGATGCCACAGATGA 1631
299 nTrpTrpProValThr.....LeuProThrGlySerGlyGlySerGlyG 314
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1632 ARGTTTCTGCTCAGCTCCCATATACCATTA..... 1662
314 lyMetGlySerGlyGlyGlyGlySerGlyGlySerGlyGlyLys 330
1662 ..... 1662
331 GlyLysGlyGlyCysPro...AlaValArgLeuLysAlaArgTrpGlnTh 346
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1663 ...AAAGTATTGAACAGGCTCCCTCGGTCTTCGACGACATCTAT 1709
346 rMetSerIleLeuProMetGluLeuTrpLysGluPheAlaGluTrpValT 363
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1710 GGAAATAATCATGCCAGAGAGAGTACAGTGAATTTAAAGACGCTTATAC 1759
363 hrAsnHisTrpArgMetLeuLysAlaValLeuGluProAlaLeuAsnVal 379
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1760 TGCAAAAGAACTTCATGTAGTCTATGCTTATCA.....CATGTA 1800
380 LysGlyLysGluGlu.....ValAlaSerAlaLeuValHisIleLeuG1 394
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1801 TGTGACAAAGACCGCAACTACTGCGCAGCATCTACTAGGATTTTCT 1850
394 nSerThrGlyLysAlaLysAspPheLeuSerAspMetAlaMetSerGluV 411
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1851 TCACGAAAAAGCTGAATCGTTGTTTATGCACACTAAATGACAGAGAAA 1900
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
411 alaAspArgPheMetGluGlu...HisLeuIlePheArgGluAsnThr 426
|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

1901 TAAGC.....ATGGAAGATGAAGCCACTACCTATTTCGAGCCACAACA 1944
427 LeuAlaThrLysAlaIleGluGluTrpMetArgLeuIleGlyGlnLysTy 443
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1945 CTTGCAAGCACCTTGATGAGCAGTATATGAAGCCACTGCTACACAGTT 1994
443 rLeuLysAspAlaIleGlyGluPheIleArgAlaLeuTrpGluSerGluG 460
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1995 TGTTCATCATGCTTTGAAGACTCTATTTTAAAGATATATGGAAGCAAGC 2044
460 luAsnCysGluValAspProIleLysCysThrAlaSerSerLeuAlaGlu 476
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2045 AGTCTTGTGAGTTAAGTCCATCAAGATTAGAAAAAATGAAGATGTGAAC 2094
477 HisGlnAlaAsn.....LeuArgMetCysGlyLeuLeuAlaLeuCysLy 491
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2095 ACTAAATTAACACACCTATTGAACATACTTTCAGAGCTT...GTGGAGAA 2141
491 sValValAsnSerHisCysValPheProArgGluLeuLysGluValPheA 508
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2142 AATATTTCATGGCTTCAGAAATACTTCCACCGACATTGAGATATATTATG 2191
508 laSerTrpArgLeuArgCysAlaGluArgGly..... 518
::: |||:|||||:|||||:|||||:|||||:|||||:|||||:
2192 GG.....TGTTTACAGAAATCTGTTACGATAGTGCCT 2226
519 ...ArgGluAspIleAlaAspArgLeuIleSerAlaSerLeuPheLeuAr 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2227 ACAATACCACCATGAGAACAGAGTTGTTAGTGGTTTGTGTTTCTTCG 2276
534 gPheLeuCysProAlaIleMetSerProSerLeuPheGlyLeuMetGlnG 551
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2277 ACPCATCTCTCTGCTGCTGATCCAGATGTTCAATATCATCTCAG 2326
551 luTrpProAspGluGlnThrSerArgThrLeuThrLeuIleAlaLysVal 567
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2327 ATTCTCCATCTCTCTGCTGCTGACAGACACTGATATTAGTGGCTAAATCT 2376
568 IleGlnAsnLeuAlaAsnPheSerLysPheThrSerLysGluAspPheLe 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2377 GTGCAGAACTTAGCAAACTTGTGGAAATTTGGAGCTAAGGAGCCCTACAT 2426
584 uGlyPheMetAsnGluPheLeuGluLeuTrpGlySerMetGlnGlnP 601
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2427 GGAAGTGTCAATCCATTCATCAAAAGCAACAACATCGTATGATCATGT 2476
601 heLeuTrpGluIleSerAsnLeuAspThrLeuThrAsnSerSerPhe 617
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2477 TTTAGATGAATTTGGAAATGTACTGAACTTCGGGACACTACAGAG... 2523
618 GluGlyTrpIleAspLeuGlyArgGluLeuSerThrLeuHisAlaLeuLe 634
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2524 CATCTAGAACCGACCTGTCCCGTGATTACAGCATTCGATGAGATTG 2573
634 u.....TrpGluValLeuProGlnLeuSerLysGlu..... 644
2574 CBTGGCTCATTCAGATGAACCTCGAACGCTCAGTAATGAGCGTGGTGAC 2623
645 .....AlaLeuLeuLysLeuGlyProLeuProArgLeuLeuSerAsp 658
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2624 AGCAGCACCTATTGAAAGCTTCTGGCTATACACAGAACTGCTTCACAA 2673
659 IleSerThrAlaLeuArgAsnProAsn...IleGlnArgGlnProSerAr 674
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2674 AACAAACACCATATACAAACCAATGATGTAGGTAGCAGCGCTTCGCC 2723
674 gGln 675
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2724 CCAG 2727

```

seq_name: /cgn2_6/ptodata/2/ina/5a_comb.seq:US-08-190-687B-7

seq_documentation_block: